(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

B4

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
17 January 2002 (17.01.2002)

PCT

(10) International Publication Number
WO 02/04520 A2

(51) International Patent Classification[7]: C07K 14/705

(21) International Application Number: PCT/US01/21448

(22) International Filing Date: 5 July 2001 (05.07.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/216,547	7 July 2000 (07.07.2000)	US
60/218,232	14 July 2000 (14.07.2000)	US
60/220,112	21 July 2000 (21.07.2000)	US
60/221,839	28 July 2000 (28.07.2000)	US

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **RAUMANN, Brigitte, E.** [US/US]; 5801 South Dorchester Avenue #3B, Chicago, IL 60637 (US). **THORNTON, Michael** [US/US]; 9 Medway Road, Woodside, CA 94062 (US). **DING, Li** [CN/US]; 3353 Alma Street #146, Palo Alto, CA 94306 (US). **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **TANG, Y., Tom** [US/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **HARLAND, Lee** [GB/GB]; 18 Chaucer Court, New Dover Road, Canterbury, Kent CT1 3AU (GB). **BURFORD, Neil** [GB/US]; 105 Wildwood Circle, Durham, CT 06422 (US). **GREENE, Barrie, D.** [US/US]; 1332 10th Avenue #104, San Franciso, CA 94122 (US). **SANJANWALA, Madhu, S.** [US/US]; 210 Sylvia Court, Los Altos, CA 94024 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **YAO, Monique, G.** [US/US]; 111 Frederick Court, Mountain View, CA 94043 (US). **YANG, Junming** [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). **PATTERSON, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US). **GANDHI, Ameena, R.** [US/US]; 837 Roble Avenue #1, Menlo Park, CA 94025 (US). **HAFALIA, April, J., A.** [US/US]; 2227 Calle de Primavera, Santa Clara, CA 95054 (US). **TRIBOULEY, Catherine, M.** [FR/US]; 1121 Tennessee Street #5, San Francisco, CA 94107 (US). **WALIA, Narinder, K.** [US/US]; 890 Davis Street #205, San Leandro, CA 94577 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **WALSH, Roderick, T.** [IE/GB]; 8 Boundary Court, St. Lawrence Road, Canterbury, Kent CT1 3EZ (GB). **RAMKUMAR, Jayalaxmi** [IN/US]; 34359 Maybird Circle, Fremont, CA 94555 (US). **LU, Yan** [CN/US]; 3885 Corrina Way, Palo Alto, CA 94303 (US). **LU, Dyung, Aina, M.** [US/US]; 233 Coy Drive, San Jose, CA 95123 (US). **AZIMZAI, Yalda** [US/US]; 5518 Boulder Canyon Drive, Castro Valley, CA 94552 (US). **LAL, Preeti** [IN/US]; P.O. Box 5142, Santa Clara, CA 95056 (US). **ELLIOTT, Vicki, S.** [US/US]; 3770 Polton Place Way, San Jose, CA 95121 (US). **NGUYEN, Danniel, B.** [US/US]; 1403 Ridgewood Drive, San Jose, CA 95118 (US). **XU, Yuming** [US/US]; 1739 Walnut Drive, Mountain View, CA 94040 (US). **SEILHAMER, Jeffrey, J.** [US/US]; 12555 La Cresta, Los Altos Hills, CA 94022 (US). **BOROWSKY, Mark, L.** [US/US]; 122 Orchard Avenue, Redwood City, CA 94061 (US). **KHAN, Farrah, A.** [IN/US]; 3617 Central Road #102, Glenview, IL 60025 (US). **KEARNEY, Liam** [US/US]; 50 Woodside Avenue, San Francisco, CA 94127 (US). **THANGAVELU, Kavitha** [IN/US]; 1950 Montecito Avenue #23, Mountain View, CA 94043 (US). **DAS, Debopriya** [IN/US]; Apartment 13, 1179 Bonita Avenue, Mountain View, CA 94040 (US). **POLICKY, Jennifer, L.** [US/US]; 1511 Jarvis Court, San Jose, CA 95118 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: TRANSPORTERS AND ION CHANNELS

(57) Abstract: The invention provides human transporters and ion channels (TRICH) and polynucleotides which identify and encode TRICH. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with aberrant expression of TRICH.

WO 02/04520 A2

TRANSPORTERS AND ION CHANNELS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of transporters and ion channels and to the use of these sequences in the diagnosis, treatment, and prevention of transport, neurological, muscle, immunological, and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

BACKGROUND OF THE INVENTION

Eukaryotic cells are surrounded and subdivided into functionally distinct organelles by hydrophobic lipid bilayer membranes which are highly impermeable to most polar molecules. Cells and organelles require transport proteins to import and export essential nutrients and metal ions including K^+, NH_4^+, P_i, SO_4^{2-}, sugars, and vitamins, as well as various metabolic waste products. Transport proteins also play roles in antibiotic resistance, toxin secretion, ion balance, synaptic neurotransmission, kidney function, intestinal absorption, tumor growth, and other diverse cell functions (Griffith, J. and C. Sansom (1998) The Transporter Facts Book, Academic Press, San Diego CA, pp. 3-29). Transport can occur by a passive concentration-dependent mechanism, or can be linked to an energy source such as ATP hydrolysis or an ion gradient. Proteins that function in transport include carrier proteins, which bind to a specific solute and undergo a conformational change that translocates the bound solute across the membrane, and channel proteins, which form hydrophilic pores that allow specific solutes to diffuse through the membrane down an electrochemical solute gradient.

Carrier proteins which transport a single solute from one side of the membrane to the other are called uniporters. In contrast, coupled transporters link the transfer of one solute with simultaneous or sequential transfer of a second solute, either in the same direction (symport) or in the opposite direction (antiport). For example, intestinal and kidney epithelium contains a variety of symporter systems driven by the sodium gradient that exists across the plasma membrane. Sodium moves into the cell down its electrochemical gradient and brings the solute into the cell with it. The sodium gradient that provides the driving force for solute uptake is maintained by the ubiquitous Na^+/K^+ ATPase system. Sodium-coupled transporters include the mammalian glucose transporter (SGLT1), iodide transporter (NIS), and multivitamin transporter (SMVT). All three transporters have twelve putative transmembrane segments, extracellular glycosylation sites, and cytoplasmically-oriented N- and C-termini. NIS plays a crucial role in the evaluation, diagnosis, and treatment of various thyroid pathologies because it is the molecular basis for radioiodide thyroid-imaging

techniques and for specific targeting of radioisotopes to the thyroid gland (Levy, O. et al. (1997) Proc. Natl. Acad. Sci. USA 94:5568-5573). SMVT is expressed in the intestinal mucosa, kidney, and placenta, and is implicated in the transport of the water-soluble vitamins, e.g., biotin and pantothenate (Prasad, P.D. et al. (1998) J. Biol. Chem. 273:7501-7506).

One of the largest families of transporters is the major facilitator superfamily (MFS), also called the uniporter-symporter-antiporter family. MFS transporters are single polypeptide carriers that transport small solutes in response to ion gradients. Members of the MFS are found in all classes of living organisms, and include transporters for sugars, oligosaccharides, phosphates, nitrates, nucleosides, monocarboxylates, and drugs. MFS transporters found in eukaryotes all have a structure comprising 12 transmembrane segments (Pao, S.S. et al. (1998) Microbiol. Molec. Biol. Rev. 62:1-34). The largest family of MFS transporters is the sugar transporter family, which includes the seven glucose transporters (GLUT1-GLUT7) found in humans that are required for the transport of glucose and other hexose sugars. These glucose transport proteins have unique tissue distributions and physiological functions. GLUT1 provides many cell types with their basal glucose requirements and transports glucose across epithelial and endothelial barrier tissues; GLUT2 facilitates glucose uptake or efflux from the liver; GLUT3 regulates glucose supply to neurons; GLUT4 is responsible for insulin-regulated glucose disposal; and GLUT5 regulates fructose uptake into skeletal muscle. Defects in glucose transporters are involved in a recently identified neurological syndrome causing infantile seizures and developmental delay, as well as glycogen storage disease, Fanconi-Bickel syndrome, and non-insulin-dependent diabetes mellitus (Mueckler, M. (1994) Eur. J. Biochem. 219:713-725; Longo, N. and L.J. Elsas (1998) Adv. Pediatr. 45:293-313).

Monocarboxylate anion transporters are proton-coupled symporters with a broad substrate specificity that includes L-lactate, pyruvate, and the ketone bodies acetate, acetoacetate, and beta-hydroxybutyrate. At least seven isoforms have been identified to date. The isoforms are predicted to have twelve transmembrane (TM) helical domains with a large intracellular loop between TM6 and TM7, and play a critical role in maintaining intracellular pH by removing the protons that are produced stoichiometrically with lactate during glycolysis. The best characterized H^+-monocarboxylate transporter is that of the erythrocyte membrane, which transports L-lactate and a wide range of other aliphatic monocarboxylates. Other cells possess H^+-linked monocarboxylate transporters with differing substrate and inhibitor selectivities. In particular, cardiac muscle and tumor cells have transporters that differ in their K_m values for certain substrates, including stereoselectivity for L- over D-lactate, and in their sensitivity to inhibitors. There are Na^+-monocarboxylate cotransporters on the luminal surface of intestinal and kidney epithelia, which allow the uptake of lactate, pyruvate, and ketone bodies in these tissues. In addition, there are specific and selective transporters for organic cations and organic anions

in organs including the kidney, intestine and liver. Organic anion transporters are selective for hydrophobic, charged molecules with electron-attracting side groups. Organic cation transporters, such as the ammonium transporter, mediate the secretion of a variety of drugs and endogenous metabolites, and contribute to the maintenance of intercellular pH (Poole, R.C. and A.P. Halestrap (1993) Am. J. Physiol. 264:C761-C782; Price, N.T. et al. (1998) Biochem. J. 329:321-328; and Martinelle, K. and I. Haggstrom (1993) J. Biotechnol. 30:339-350).

ATP-binding cassette (ABC) transporters are members of a superfamily of membrane proteins that transport substances ranging from small molecules such as ions, sugars, amino acids, peptides, and phospholipids, to lipopeptides, large proteins, and complex hydrophobic drugs. ABC transporters consist of four modules: two nucleotide-binding domains (NBD), which hydrolyze ATP to supply the energy required for transport, and two membrane-spanning domains (MSD), each containing six putative transmembrane segments. These four modules may be encoded by a single gene, as is the case for the cystic fibrosis transmembrane regulator (CFTR), or by separate genes. When encoded by separate genes, each gene product contains a single NBD and MSD. These "half-molecules" form homo- and heterodimers, such as Tap1 and Tap2, the endoplasmic reticulum-based major histocompatibility (MHC) peptide transport system. Several genetic diseases are attributed to defects in ABC transporters, such as the following diseases and their corresponding proteins: cystic fibrosis (CFTR, an ion channel), adrenoleukodystrophy (adrenoleukodystrophy protein, ALDP), Zellweger syndrome (peroxisomal membrane protein-70, PMP70), and hyperinsulinemic hypoglycemia (sulfonylurea receptor, SUR). Overexpression of the multidrug resistance (MDR) protein, another ABC transporter, in human cancer cells makes the cells resistant to a variety of cytotoxic drugs used in chemotherapy (Taglicht, D. and S. Michaelis (1998) Meth. Enzymol. 292:130-162).

A number of metal ions such as iron, zinc, copper, cobalt, manganese, molybdenum, selenium, nickel, and chromium are important as cofactors for a number of enzymes. For example, copper is involved in hemoglobin synthesis, connective tissue metabolism, and bone development, by acting as a cofactor in oxidoreductases such as superoxide dismutase, ferroxidase (ceruloplasmin), and lysyl oxidase. Copper and other metal ions must be provided in the diet, and are absorbed by transporters in the gastrointestinal tract. Plasma proteins transport the metal ions to the liver and other target organs, where specific transporters move the ions into cells and cellular organelles as needed. Imbalances in metal ion metabolism have been associated with a number of disease states (Danks, D.M. (1986) J. Med. Genet. 23:99-106).

Transport of fatty acids across the plasma membrane can occur by diffusion, a high capacity, low affinity process. However, under normal physiological conditions a significant fraction of fatty acid transport appears to occur via a high affinity, low capacity protein-mediated transport process.

Fatty acid transport protein (FATP), an integral membrane protein with four transmembrane segments, is expressed in tissues exhibiting high levels of plasma membrane fatty acid flux, such as muscle, heart, and adipose. Expression of FATP is upregulated in 3T3-L1 cells during adipose conversion, and expression in COS7 fibroblasts elevates uptake of long-chain fatty acids (Hui, T.Y. et al. (1998) J. Biol. Chem. 273:27420-27429).

Mitochondrial carrier proteins are transmembrane-spanning proteins which transport ions and charged metabolites between the cytosol and the mitochondrial matrix. Examples include the ADP, ATP carrier protein; the 2-oxoglutarate/malate carrier; the phosphate carrier protein; the pyruvate carrier; the dicarboxylate carrier which transports malate, succinate, fumarate, and phosphate; the tricarboxylate carrier which transports citrate and malate; and the Grave's disease carrier protein, a protein recognized by IgG in patients with active Grave's disease, an autoimmune disorder resulting in hyperthyroidism. Proteins in this family consist of three tandem repeats of an approximately 100 amino acid domain, each of which contains two transmembrane regions (Stryer, L. (1995) Biochemistry, W.H. Freeman and Company, New York NY, p. 551; PROSITE PDOC00189 Mitochondrial energy transfer proteins signature; Online Mendelian Inheritance in Man (OMIM) *275000 Graves Disease).

This class of transporters also includes the mitochondrial uncoupling proteins, which create proton leaks across the inner mitochondrial membrane, thus uncoupling oxidative phosphorylation from ATP synthesis. The result is energy dissipation in the form of heat. Mitochondrial uncoupling proteins have been implicated as modulators of thermoregulation and metabolic rate, and have been proposed as potential targets for drugs against metabolic diseases such as obesity (Ricquier, D. et al. (1999) J. Int. Med. 245:637-642).

Ion Channels

The electrical potential of a cell is generated and maintained by controlling the movement of ions across the plasma membrane. The movement of ions requires ion channels, which form ion-selective pores within the membrane. There are two basic types of ion channels, ion transporters and gated ion channels. Ion transporters utilize the energy obtained from ATP hydrolysis to actively transport an ion against the ion's concentration gradient. Gated ion channels allow passive flow of an ion down the ion's electrochemical gradient under restricted conditions. Together, these types of ion channels generate, maintain, and utilize an electrochemical gradient that is used in 1) electrical impulse conduction down the axon of a nerve cell, 2) transport of molecules into cells against concentration gradients, 3) initiation of muscle contraction, and 4) endocrine cell secretion.

Ion Transporters

Ion transporters generate and maintain the resting electrical potential of a cell. Utilizing the energy derived from ATP hydrolysis, they transport ions against the ion's concentration gradient.

These transmembrane ATPases are divided into three families. The phosphorylated (P) class ion transporters, including Na^+-K^+ ATPase, Ca^{2+}-ATPase, and H^+-ATPase, are activated by a phosphorylation event. P-class ion transporters are responsible for maintaining resting potential distributions such that cytosolic concentrations of Na^+ and Ca^{2+} are low and cytosolic concentration of K^+ is high. The vacuolar (V) class of ion transporters includes H^+ pumps on intracellular organelles, such as lysosomes and Golgi. V-class ion transporters are responsible for generating the low pH within the lumen of these organelles that is required for function. The coupling factor (F) class consists of H^+ pumps in the mitochondria. F-class ion transporters utilize a proton gradient to generate ATP from ADP and inorganic phosphate (P_i).

The P-ATPases are hexamers of a 100 kD subunit with ten transmembrane domains and several large cytoplasmic regions that may play a role in ion binding (Scarborough, G.A. (1999) Curr. Opin. Cell Biol. 11:517-522). The V-ATPases are composed of two functional domains: the V_1 domain, a peripheral complex responsible for ATP hydrolysis; and the V_0 domain, an integral complex responsible for proton translocation across the membrane. The F-ATPases are structurally and evolutionarily related to the V-ATPases. The F-ATPase F_0 domain contains 12 copies of the c subunit, a highly hydrophobic protein composed of two transmembrane domains and containing a single buried carboxyl group in TM2 that is essential for proton transport. The V-ATPase V_0 domain contains three types of homologous c subunits with four or five transmembrane domains and the essential carboxyl group in TM4 or TM3. Both types of complex also contain a single a subunit that may be involved in regulating the pH dependence of activity (Forgac, M. (1999) J. Biol. Chem. 274:12951-12954).

The resting potential of the cell is utilized in many processes involving carrier proteins and gated ion channels. Carrier proteins utilize the resting potential to transport molecules into and out of the cell. Amino acid and glucose transport into many cells is linked to sodium ion co-transport (symport) so that the movement of Na^+ down an electrochemical gradient drives transport of the other molecule up a concentration gradient. Similarly, cardiac muscle links transfer of Ca^{2+} out of the cell with transport of Na^+ into the cell (antiport).

Gated Ion Channels

Gated ion channels control ion flow by regulating the opening and closing of pores. The ability to control ion flux through various gating mechanisms allows ion channels to mediate such diverse signaling and homeostatic functions as neuronal and endocrine signaling, muscle contraction, fertilization, and regulation of ion and pH balance. Gated ion channels are categorized according to the manner of regulating the gating function. Mechanically-gated channels open their pores in response to mechanical stress; voltage-gated channels (e.g., Na^+, K^+, Ca^{2+}, and Cl^- channels) open their pores in response to changes in membrane potential; and ligand-gated channels (e.g.,

acetylcholine-, serotonin-, and glutamate-gated cation channels, and GABA- and glycine-gated chloride channels) open their pores in the presence of a specific ion, nucleotide, or neurotransmitter. The gating properties of a particular ion channel (i.e., its threshold for and duration of opening and closing) are sometimes modulated by association with auxiliary channel proteins and/or post translational modifications, such as phosphorylation.

Mechanically-gated or mechanosensitive ion channels act as transducers for the senses of touch, hearing, and balance, and also play important roles in cell volume regulation, smooth muscle contraction, and cardiac rhythm generation. A stretch-inactivated channel (SIC) was recently cloned from rat kidney. The SIC channel belongs to a group of channels which are activated by pressure or stress on the cell membrane and conduct both Ca^{2+} and Na^{+} (Suzuki, M. et al. (1999) J. Biol. Chem. 274:6330-6335).

The pore-forming subunits of the voltage-gated cation channels form a superfamily of ion channel proteins. The characteristic domain of these channel proteins comprises six transmembrane domains (S1-S6), a pore-forming region (P) located between S5 and S6, and intracellular amino and carboxy termini. In the Na^{+} and Ca^{2+} subfamilies, this domain is repeated four times, while in the K^{+} channel subfamily, each channel is formed from a tetramer of either identical or dissimilar subunits. The P region contains information specifying the ion selectivity for the channel. In the case of K^{+} channels, a GYG tripeptide is involved in this selectivity (Ishii, T.M. et al. (1997) Proc. Natl. Acad. Sci. USA 94:11651-11656).

Voltage-gated Na^{+} and K^{+} channels are necessary for the function of electrically excitable cells, such as nerve and muscle cells. Action potentials, which lead to neurotransmitter release and muscle contraction, arise from large, transient changes in the permeability of the membrane to Na^{+} and K^{+} ions. Depolarization of the membrane beyond the threshold level opens voltage-gated Na^{+} channels. Sodium ions flow into the cell, further depolarizing the membrane and opening more voltage-gated Na^{+} channels, which propagates the depolarization down the length of the cell. Depolarization also opens voltage-gated potassium channels. Consequently, potassium ions flow outward, which leads to repolarization of the membrane. Voltage-gated channels utilize charged residues in the fourth transmembrane segment (S4) to sense voltage change. The open state lasts only about 1 millisecond, at which time the channel spontaneously converts into an inactive state that cannot be opened irrespective of the membrane potential. Inactivation is mediated by the channel's N-terminus, which acts as a plug that closes the pore. The transition from an inactive to a closed state requires a return to resting potential.

Voltage-gated Na^{+} channels are heterotrimeric complexes composed of a 260 kDa pore-forming α subunit that associates with two smaller auxiliary subunits, $\beta 1$ and $\beta 2$. The $\beta 2$ subunit is a integral

membrane glycoprotein that contains an extracellular Ig domain, and its association with α and β1 subunits correlates with increased functional expression of the channel, a change in its gating properties, as well as an increase in whole cell capacitance due to an increase in membrane surface area (Isom, L.L. et al. (1995) Cell 83:433-442).

Non voltage-gated Na^+ channels include the members of the amiloride-sensitive Na^+ channel/degenerin (NaC/DEG) family. Channel subunits of this family are thought to consist of two transmembrane domains flanking a long extracellular loop, with the amino and carboxyl termini located within the cell. The NaC/DEG family includes the epithelial Na^+ channel (ENaC) involved in Na^+ reabsorption in epithelia including the airway, distal colon, cortical collecting duct of the kidney, and exocrine duct glands. Mutations in ENaC result in pseudohypoaldosteronism type 1 and Liddle's syndrome (pseudohyperaldosteronism). The NaC/DEG family also includes the recently characterized H^+-gated cation channels or acid-sensing ion channels (ASIC). ASIC subunits are expressed in the brain and form heteromultimeric Na^+-permeable channels. These channels require acid pH fluctuations for activation. ASIC subunits show homology to the degenerins, a family of mechanically-gated channels originally isolated from C. elegans. Mutations in the degenerins cause neurodegeneration. ASIC subunits may also have a role in neuronal function, or in pain perception, since tissue acidosis causes pain (Waldmann, R. and M. Lazdunski (1998) Curr. Opin. Neurobiol. 8:418-424; Eglen, R.M. et al. (1999) Trends Pharmacol. Sci. 20:337-342).

K^+ channels are located in all cell types, and may be regulated by voltage, ATP concentration, or second messengers such as Ca^{2+} and cAMP. In non-excitable tissue, K^+ channels are involved in protein synthesis, control of endocrine secretions, and the maintenance of osmotic equilibrium across membranes. In neurons and other excitable cells, in addition to regulating action potentials and repolarizing membranes, K^+ channels are responsible for setting resting membrane potential. The cytosol contains non-diffusible anions and, to balance this net negative charge, the cell contains a Na^+-K^+ pump and ion channels that provide the redistribution of Na^+, K^+, and Cl^-. The pump actively transports Na^+ out of the cell and K^+ into the cell in a 3:2 ratio. Ion channels in the plasma membrane allow K^+ and Cl^- to flow by passive diffusion. Because of the high negative charge within the cytosol, Cl^- flows out of the cell. The flow of K^+ is balanced by an electromotive force pulling K^+ into the cell, and a K^+ concentration gradient pushing K^+ out of the cell. Thus, the resting membrane potential is primarily regulated by K^+ flow (Salkoff, L. and T. Jegla (1995) Neuron 15:489-492).

Potassium channel subunits of the Shaker-like superfamily all have the characteristic six transmembrane/1 pore domain structure. Four subunits combine as homo- or heterotetramers to form functional K channels. These pore-forming subunits also associate with various cytoplasmic β subunits that alter channel inactivation kinetics. The Shaker-like channel family includes the voltage-

gated K^+ channels as well as the delayed rectifier type channels such as the human ether-a-go-go related gene (HERG) associated with long QT, a cardiac dysrythmia syndrome (Curran, M.E. (1998) Curr. Opin. Biotechnol. 9:565-572; Kaczorowski, G.J. and M.L. Garcia (1999) Curr. Opin. Chem. Biol. 3:448-458).

A second superfamily of K^+ channels is composed of the inward rectifying channels (Kir). Kir channels have the property of preferentially conducting K^+ currents in the inward direction. These proteins consist of a single potassium selective pore domain and two transmembrane domains, which correspond to the fifth and sixth transmembrane domains of voltage-gated K^+ channels. Kir subunits also associate as tetramers. The Kir family includes ROMK1, mutations in which lead to Bartter syndrome, a renal tubular disorder. Kir channels are also involved in regulation of cardiac pacemaker activity, seizures and epilepsy, and insulin regulation (Doupnik, C.A. et al. (1995) Curr. Opin. Neurobiol. 5:268-277; Curran, supra).

The recently recognized TWIK K^+ channel family includes the mammalian TWIK-1, TREK-1 and TASK proteins. Members of this family possess an overall structure with four transmembrane domains and two P domains. These proteins are probably involved in controlling the resting potential in a large set of cell types (Duprat, F. et al. (1997) EMBO J 16:5464-5471).

The voltage-gated Ca^{2+} channels have been classified into several subtypes based upon their electrophysiological and pharmacological characteristics. L-type Ca^{2+} channels are predominantly expressed in heart and skeletal muscle where they play an essential role in excitation-contraction coupling. T-type channels are important for cardiac pacemaker activity, while N-type and P/Q-type channels are involved in the control of neurotransmitter release in the central and peripheral nervous system. The L-type and N-type voltage-gated Ca^{2+} channels have been purified and, though their functions differ dramatically, they have similar subunit compositions. The channels are composed of three subunits. The α_1 subunit forms the membrane pore and voltage sensor, while the $\alpha_2\delta$ and β subunits modulate the voltage-dependence, gating properties, and the current amplitude of the channel. These subunits are encoded by at least six α_1, one $\alpha_2\delta$, and four β genes. A fourth subunit, γ, has been identified in skeletal muscle (Walker, D. et al. (1998) J. Biol. Chem. 273:2361-2367; McCleskey, E.W. (1994) Curr. Opin. Neurobiol. 4:304-312).

The transient receptor family (Trp) of calcium ion channels are thought to mediate capacitative calcium entry (CCE). CCE is the Ca^{2+} influx into cells to resupply Ca^{2+} stores depleted by the action of inositol triphosphate (IP3) and other agents in response to numerous hormones and growth factors. Trp and Trp-like were first cloned from Drosophila and have similarity to voltage gated Ca2+ channels in the S3 through S6 regions. This suggests that Trp and/or related proteins may form mammalian CCC entry channels (Zhu, X. et al. (1996) Cell 85:661-671; Boulay, G. et al. (1997) J. Biol. Chem.

272:29672-29680). Melastatin is a gene isolated in both the mouse and human, and whose expression in melanoma cells is inversely correlated with melanoma aggressiveness in vivo. The human cDNA transcript corresponds to a 1533-amino acid protein having homology to members of the Trp family. It has been proposed that the combined use of malastatin mRNA expression status and tumor thickness might allow for the determination of subgroups of patients at both low and high risk for developing metastatic disease (Duncan, L.M. et al (2001) J. Clin. Oncol. 19:568-576).

Chloride channels are necessary in endocrine secretion and in regulation of cytosolic and organelle pH. In secretory epithelial cells, Cl^- enters the cell across a basolateral membrane through an Na^+, K^+/Cl^- cotransporter, accumulating in the cell above its electrochemical equilibrium concentration. Secretion of Cl^- from the apical surface, in response to hormonal stimulation, leads to flow of Na^+ and water into the secretory lumen. The cystic fibrosis transmembrane conductance regulator (CFTR) is a chloride channel encoded by the gene for cystic fibrosis, a common fatal genetic disorder in humans. CFTR is a member of the ABC transporter family, and is composed of two domains each consisting of six transmembrane domains followed by a nucleotide-binding site. Loss of CFTR function decreases transepithelial water secretion and, as a result, the layers of mucus that coat the respiratory tree, pancreatic ducts, and intestine are dehydrated and difficult to clear. The resulting blockage of these sites leads to pancreatic insufficiency, "meconium ileus", and devastating "chronic obstructive pulmonary disease" (Al-Awqati, Q. et al. (1992) J. Exp. Biol. 172:245-266).

The voltage-gated chloride channels (CLC) are characterized by 10-12 transmembrane domains, as well as two small globular domains known as CBS domains. The CLC subunits probably function as homotetramers. CLC proteins are involved in regulation of cell volume, membrane potential stabilization, signal transduction, and transepithelial transport. Mutations in CLC-1, expressed predominantly in skeletal muscle, are responsible for autosomal recessive generalized myotonia and autosomal dominant myotonia congenita, while mutations in the kidney channel CLC-5 lead to kidney stones (Jentsch, T.J. (1996) Curr. Opin. Neurobiol. 6:303-310).

Ligand-gated channels open their pores when an extracellular or intracellular mediator binds to the channel. Neurotransmitter-gated channels are channels that open when a neurotransmitter binds to their extracellular domain. These channels exist in the postsynaptic membrane of nerve or muscle cells. There are two types of neurotransmitter-gated channels. Sodium channels open in response to excitatory neurotransmitters, such as acetylcholine, glutamate, and serotonin. This opening causes an influx of Na^+ and produces the initial localized depolarization that activates the voltage-gated channels and starts the action potential. Chloride channels open in response to inhibitory neurotransmitters, such as γ-aminobutyric acid (GABA) and glycine, leading to hyperpolarization of the membrane and the subsequent generation of an action potential. Neurotransmitter-gated ion channels have four

transmembrane domains and probably function as pentamers (Jentsch, supra). Amino acids in the second transmembrane domain appear to be important in determining channel permeation and selectivity (Sather, W.A. et al. (1994) Curr. Opin. Neurobiol. 4:313-323).

Ligand-gated channels can be regulated by intracellular second messengers. For example, calcium-activated K^+ channels are gated by internal calcium ions. In nerve cells, an influx of calcium during depolarization opens K^+ channels to modulate the magnitude of the action potential (Ishi et al., supra). The large conductance (BK) channel has been purified from brain and its subunit composition determined. The α subunit of the BK channel has seven rather than six transmembrane domains in contrast to voltage-gated K^+ channels. The extra transmembrane domain is located at the subunit N-terminus. A 28-amino-acid stretch in the C-terminal region of the subunit (the "calcium bowl" region) contains many negatively charged residues and is thought to be the region responsible for calcium binding. The β subunit consists of two transmembrane domains connected by a glycosylated extracellular loop, with intracellular N- and C-termini (Kaczorowski, supra; Vergara, C. et al. (1998) Curr. Opin. Neurobiol. 8:321-329).

Cyclic nucleotide-gated (CNG) channels are gated by cytosolic cyclic nucleotides. The best examples of these are the cAMP-gated Na^+ channels involved in olfaction and the cGMP-gated cation channels involved in vision. Both systems involve ligand-mediated activation of a G-protein coupled receptor which then alters the level of cyclic nucleotide within the cell. CNG channels also represent a major pathway for Ca^{2+} entry into neurons, and play roles in neuronal development and plasticity. CNG channels are tetramers containing at least two types of subunits, an α subunit which can form functional homomeric channels, and a β subunit, which modulates the channel properties. All CNG subunits have six transmembrane domains and a pore forming region between the fifth and sixth transmembrane domains, similar to voltage-gated K^+ channels. A large C-terminal domain contains a cyclic nucleotide binding domain, while the N-terminal domain confers variation among channel subtypes (Zufall, F. et al. (1997) Curr. Opin. Neurobiol. 7:404-412).

The activity of other types of ion channel proteins may also be modulated by a variety of intracellular signalling proteins. Many channels have sites for phosphorylation by one or more protein kinases including protein kinase A, protein kinase C, tyrosine kinase, and casein kinase II, all of which regulate ion channel activity in cells. Kir channels are activated by the binding of the Gβγ subunits of heterotrimeric G-proteins (Reimann, F. and F.M. Ashcroft (1999) Curr. Opin. Cell. Biol. 11:503-508). Other proteins are involved in the localization of ion channels to specific sites in the cell membrane. Such proteins include the PDZ domain proteins known as MAGUKs (membrane-associated guanylate kinases) which regulate the clustering of ion channels at neuronal synapses (Craven, S.E. and D.S. Bredt (1998) Cell 93:495-498).

Disease Correlation

The etiology of numerous human diseases and disorders can be attributed to defects in the transport of molecules across membranes. Defects in the trafficking of membrane-bound transporters and ion channels are associated with several disorders, e.g., cystic fibrosis, glucose-galactose malabsorption syndrome, hypercholesterolemia, von Gierke disease, and certain forms of diabetes mellitus. Single-gene defect diseases resulting in an inability to transport small molecules across membranes include, e.g., cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease (van't Hoff, W.G. (1996) Exp. Nephrol. 4:253-262; Talente, G.M. et al. (1994) Ann. Intern. Med. 120:218-226; and Chillon, M. et al. (1995) New Engl. J. Med. 332:1475-1480).

Human diseases caused by mutations in ion channel genes include disorders of skeletal muscle, cardiac muscle, and the central nervous system. Mutations in the pore-forming subunits of sodium and chloride channels cause myotonia, a muscle disorder in which relaxation after voluntary contraction is delayed. Sodium channel myotonias have been treated with channel blockers. Mutations in muscle sodium and calcium channels cause forms of periodic paralysis, while mutations in the sarcoplasmic calcium release channel, T-tubule calcium channel, and muscle sodium channel cause malignant hyperthermia. Cardiac arrythmia disorders such as the long QT syndromes and idiopathic ventricular fibrillation are caused by mutations in potassium and sodium channels (Cooper, E.C. and L.Y. Jan (1998) Proc. Natl. Acad. Sci. USA 96:4759-4766). All four known human idiopathic epilepsy genes code for ion channel proteins (Berkovic, S.F. and I.E. Scheffer (1999) Curr. Opin. Neurology 12:177-182). Other neurological disorders such as ataxias, hemiplegic migraine and hereditary deafness can also result from mutations in ion channel genes (Jen, J. (1999) Curr. Opin. Neurobiol. 9:274-280; Cooper, supra).

Ion channels have been the target for many drug therapies. Neurotransmitter-gated channels have been targeted in therapies for treatment of insomnia, anxiety, depression, and schizophrenia. Voltage-gated channels have been targeted in therapies for arrhythmia, ischemic stroke, head trauma, and neurodegenerative disease (Taylor, C.P. and L.S. Narasimhan (1997) Adv. Pharmacol. 39:47-98). Various classes of ion channels also play an important role in the perception of pain, and thus are potential targets for new analgesics. These include the vanilloid-gated ion channels, which are activated by the vanilloid capsaicin, as well as by noxious heat. Local anesthetics such as lidocaine and mexiletine which blockade voltage-gated Na^+ channels have been useful in the treatment of neuropathic pain (Eglen, supra).

Ion channels in the immune system have recently been suggested as targets for immunomodulation. T-cell activation depends upon calcium signaling, and a diverse set of T-cell specific ion channels has been characterized that affect this signaling process. Channel blocking agents

can inhibit secretion of lymphokines, cell proliferation, and killing of target cells. A peptide antagonist of the T-cell potassium channel Kv1.3 was found to suppress delayed-type hypersensitivity and allogenic responses in pigs, validating the idea of channel blockers as safe and efficacious immunosuppressants (Cahalan, M.D. and K.G. Chandy (1997) Curr. Opin. Biotechnol. 8:749-756).

The discovery of new transporters and ion channels, and the polynucleotides encoding them, satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of transport, neurological, muscle, immunological, and cell proliferative disorders, and in the assessment of the effects of exogenous compounds on the expression of nucleic acid and amino acid sequences of transporters and ion channels.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, transporters and ion channels, referred to collectively as "TRICH" and individually as "TRICH-1," "TRICH-2," "TRICH-3," "TRICH-4," "TRICH-5," "TRICH-6," "TRICH-7," "TRICH-8," "TRICH-9," "TRICH-10," "TRICH-11," "TRICH-12," "TRICH-13," "TRICH-14," "TRICH-15," "TRICH-16," "TRICH-17," "TRICH-18," "TRICH-19," "TRICH-20," "TRICH-21," "TRICH-22," "TRICH-23," "TRICH-24," "TRICH-25," "TRICH-26," "TRICH-27," "TRICH-28," "TRICH-29," "TRICH-30," "TRICH-31," and "TRICH-32." In one aspect, the invention provides an isolated polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-32.

The invention further provides an isolated polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-32. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:33-64.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

The invention further provides an isolated polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide selected from the group consisting of a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, c) a polynucleotide complementary to the polynucleotide of a), d) a polynucleotide complementary to the polynucleotide of b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and a pharmaceutically acceptable excipient. In one embodiment, the composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) combining the polypeptide with at least one test compound under

suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide selected from the group consisting of a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:33-64, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, iii) a polynucleotide having a sequence complementary to i), iv) a polynucleotide complementary to the polynucleotide of ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide selected from the group consisting of i) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, ii) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64, iii) a polynucleotide complementary to the polynucleotide of i), iv) a polynucleotide complementary to the

polynucleotide of ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of a polynucleotide sequence selected from the group consisting of i)-v) above; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the present invention.

Table 2 shows the GenBank identification number and annotation of the nearest GenBank homolog for polypeptides of the invention. The probability score for the match between each polypeptide and its GenBank homolog is also shown.

Table 3 shows structural features of polypeptide sequences of the invention, including predicted motifs and domains, along with the methods, algorithms, and searchable databases used for analysis of the polypeptides.

Table 4 lists the cDNA and/or genomic DNA fragments which were used to assemble polynucleotide sequences of the invention, along with selected fragments of the polynucleotide sequences.

Table 5 shows the representative cDNA library for polynucleotides of the invention.

Table 6 provides an appendix which describes the tissues and vectors used for construction of the cDNA libraries shown in Table 5.

Table 7 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a

reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"TRICH" refers to the amino acid sequences of substantially purified TRICH obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of TRICH. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

An "allelic variant" is an alternative form of the gene encoding TRICH. Allelic variants may result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding TRICH include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as TRICH or a polypeptide with at least one functional characteristic of TRICH. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding TRICH, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding TRICH. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent TRICH. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge,

solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of TRICH is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of TRICH. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of TRICH either by directly interacting with TRICH or by acting on components of the biological pathway in which TRICH participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind TRICH polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic TRICH, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding TRICH or fragments of TRICH may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (Applied Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the side chain.

A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

"Differential expression" refers to increased or upregulated; or decreased, downregulated, or

absent gene or protein expression, determined by comparing at least two different samples. Such comparisons may be carried out between, for example, a treated and an untreated sample, or a diseased and a normal sample.

A "fragment" is a unique portion of TRICH or the polynucleotide encoding TRICH which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50%) of a polypeptide as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:33-64 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:33-64, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:33-64 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:33-64 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:33-64 and the region of SEQ ID NO:33-64 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-32 is encoded by a fragment of SEQ ID NO:33-64. A fragment of SEQ ID NO:1-32 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-32. For example, a fragment of SEQ ID NO:1-32 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-32. The precise length of a fragment of SEQ ID NO:1-32 and the region of SEQ ID NO:1-32 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full length" polynucleotide sequence encodes a "full length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to

the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore, achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for

example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance; a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions for nucleic acid hybridization are well known and can be found in Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour.

Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 µg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of TRICH which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of TRICH which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of TRICH. For example, modulation may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of TRICH.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or

synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an TRICH may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of TRICH.

"Probe" refers to nucleic acid sequences encoding TRICH, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for example Sambrook, J. et al. (1989) Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al. (1987) Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al. (1990) PCR

Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing TRICH, nucleic acids encoding TRICH, or fragments thereof may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters, chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers,

microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed cells" includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant

identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides will generally have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool, Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 91%, at least 92%, at least 93%, at least 94%, at least 95%, at least 96%, at least 97%, at least 98%, or at least 99% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human transporters and ion channels (TRICH), the polynucleotides encoding TRICH, and the use of these compositions for the diagnosis, treatment, or prevention of transport, neurological, muscle, immunological, and cell proliferative disorders.

Table 1 summarizes the nomenclature for the full length polynucleotide and polypeptide sequences of the invention. Each polynucleotide and its corresponding polypeptide are correlated to a single Incyte project identification number (Incyte Project ID). Each polypeptide sequence is denoted by both a polypeptide sequence identification number (Polypeptide SEQ ID NO:) and an Incyte polypeptide sequence number (Incyte Polypeptide ID) as shown. Each polynucleotide sequence is denoted by both a polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and an Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) as shown.

Table 2 shows sequences with homology to the polypeptides of the invention as identified by BLAST analysis against the GenBank protein (genpept) database. Columns 1 and 2 show the polypeptide sequence identification number (Polypeptide SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for polypeptides of the invention. Column 3 shows the GenBank identification number (Genbank ID NO:) of the nearest GenBank homolog. Column 4 shows the probability score for the match between each polypeptide and its GenBank

homolog. Column 5 shows the annotation of the GenBank homolog along with relevant citations where applicable, all of which are expressly incorporated by reference herein.

Table 3 shows various structural features of the polypeptides of the invention. Columns 1 and 2 show the polypeptide sequence identification number (SEQ ID NO:) and the corresponding Incyte polypeptide sequence number (Incyte Polypeptide ID) for each polypeptide of the invention. Column 3 shows the number of amino acid residues in each polypeptide. Column 4 shows potential phosphorylation sites, and column 5 shows potential glycosylation sites, as determined by the MOTIFS program of the GCG sequence analysis software package (Genetics Computer Group, Madison WI). Column 6 shows amino acid residues comprising signature sequences, domains, and motifs. Column 7 shows analytical methods for protein structure/function analysis and in some cases, searchable databases to which the analytical methods were applied.

Together, Tables 2 and 3 summarize the properties of polypeptides of the invention, and these properties establish that the claimed polypeptides are transporters and ion channels. For example, SEQ ID NO:5 is 83% identical to rat GABA receptor rho-3 subunit precursor (GenBank ID g1060975) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 1.7e-206, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:5 also contains a neurotransmitter-gated ion channel domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:5 is a neurotransmitter-gated ion channel. In an alternate example, SEQ ID NO:16 is 57% identical to human Na+/glucose cotransporter (GenBank ID g338055) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 2.4e-181, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:16 also contains a sodium:solute symporter family domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:16 is a Na+/glucose cotransporter. In an alternate example, SEQ ID NO:27 is 53% identical to human ATP-binding cassette transporter-1 (ABC-1) (GenBank ID g4128033) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 0.0, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:27 also contains an ABC transporter domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from

BLIMPS, MOTIFS, and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:27 is an ABC transporter. In an alternate example, SEQ ID NO:12 is 45% identical to rat thyroid sodium/iodide symporter NIS (GenBank ID g1399954) as determined by the Basic Local Alignment Search Tool (BLAST). (See Table 2.) The BLAST probability score is 3.0e-143, which indicates the probability of obtaining the observed polypeptide sequence alignment by chance. SEQ ID NO:12 also contains a sodium:solute symporter family domain as determined by searching for statistically significant matches in the hidden Markov model (HMM)-based PFAM database of conserved protein family domains. (See Table 3.) Data from BLIMPS and PROFILESCAN analyses provide further corroborative evidence that SEQ ID NO:12 is a sodium:solute symporter. SEQ ID NO:1-4, SEQ ID NO:6-11, SEQ ID NO:13-15, SEQ ID NO:17-26, and SEQ ID NO:28-32 were analyzed and annotated in a similar manner. The algorithms and parameters for the analysis of SEQ ID NO:1-32 are described in Table 7.

As shown in Table 4, the full length polynucleotide sequences of the present invention were assembled using cDNA sequences or coding (exon) sequences derived from genomic DNA, or any combination of these two types of sequences. Columns 1 and 2 list the polynucleotide sequence identification number (Polynucleotide SEQ ID NO:) and the corresponding Incyte polynucleotide consensus sequence number (Incyte Polynucleotide ID) for each polynucleotide of the invention. Column 3 shows the length of each polynucleotide sequence in basepairs. Column 4 lists fragments of the polynucleotide sequences which are useful, for example, in hybridization or amplification technologies that identify SEQ ID NO:33-64 or that distinguish between SEQ ID NO:33-64 and related polynucleotide sequences. Column 5 shows identification numbers corresponding to cDNA sequences, coding sequences (exons) predicted from genomic DNA, and/or sequence assemblages comprised of both cDNA and genomic DNA. These sequences were used to assemble the full length polynucleotide sequences of the invention. Columns 6 and 7 of Table 4 show the nucleotide start (5′) and stop (3′) positions of the cDNA and/or genomic sequences in column 5 relative to their respective full length sequences.

The identification numbers in Column 5 of Table 4 may refer specifically, for example, to Incyte cDNAs along with their corresponding cDNA libraries. For example, 6724643H1 is the identification number of an Incyte cDNA sequence, and LUNLTMT01 is the cDNA library from which it is derived. Incyte cDNAs for which cDNA libraries are not indicated were derived from pooled cDNA libraries (e.g., 71495515V1). Alternatively, the identification numbers in column 5 may refer to GenBank cDNAs or ESTs (e.g., g5746200) which contributed to the assembly of the full length polynucleotide sequences. In addition, the identification numbers in column 5 may identify sequences derived from the ENSEMBL (The Sanger Centre, Cambridge, UK) database (*i.e.*, those sequences

including the designation "ENST"). Alternatively, the identification numbers in column 5 may be derived from the NCBI RefSeq Nucleotide Sequence Records Database (*i.e.*, those sequences including the designation "NM" or "NT") or the NCBI RefSeq Protein Sequence Records (*i.e.*, those sequences including the designation "NP"). Alternatively, the identification numbers in column 5 may refer to assemblages of both cDNA and Genscan-predicted exons brought together by an "exon stitching" algorithm. For example, FL_*XXXXXX*_N_1_N_2_*YYYYY*_N_3_N_4 represents a "stitched" sequence in which *XXXXXX* is the identification number of the cluster of sequences to which the algorithm was applied, and *YYYYY* is the number of the prediction generated by the algorithm, and $N_{1,2,3...}$, if present, represent specific exons that may have been manually edited during analysis (See Example V). Alternatively, the identification numbers in column 5 may refer to assemblages of exons brought together by an "exon-stretching" algorithm. For example, FL*XXXXXX*_g*AAAAA*_g*BBBBB*_1_*N* is the identification number of a "stretched" sequence, with *XXXXXX* being the Incyte project identification number, g*AAAAA* being the GenBank identification number of the human genomic sequence to which the "exon-stretching" algorithm was applied, g*BBBBB* being the GenBank identification number or NCBI RefSeq identification number of the nearest GenBank protein homolog, and *N* referring to specific exons (See Example V). In instances where a RefSeq sequence was used as a protein homolog for the "exon-stretching" algorithm, a RefSeq identifier (denoted by "NM," "NP," or "NT") may be used in place of the GenBank identifier (*i.e.*, g*BBBBB*).

Alternatively, a prefix identifies component sequences that were hand-edited, predicted from genomic DNA sequences, or derived from a combination of sequence analysis methods. The following Table lists examples of component sequence prefixes and corresponding sequence analysis methods associated with the prefixes (see Example IV and Example V).

Prefix	Type of analysis and/or examples of programs
GNN, GFG, ENST	Exon prediction from genomic sequences using, for example, GENSCAN (Stanford University, CA, USA) or FGENES (Computer Genomics Group, The Sanger Centre, Cambridge, UK).
GBI	Hand-edited analysis of genomic sequences.
FL	Stitched or stretched genomic sequences (see Example V).

In some cases, Incyte cDNA coverage redundant with the sequence coverage shown in column 5 was obtained to confirm the final consensus polynucleotide sequence, but the relevant Incyte cDNA identification numbers are not shown.

Table 5 shows the representative cDNA libraries for those full length polynucleotide sequences which were assembled using Incyte cDNA sequences. The representative cDNA library is the Incyte

cDNA library which is most frequently represented by the Incyte cDNA sequences which were used to assemble and confirm the above polynucleotide sequences. The tissues and vectors which were used to construct the cDNA libraries shown in Table 5 are described in Table 6.

The invention also encompasses TRICH variants. A preferred TRICH variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the TRICH amino acid sequence, and which contains at least one functional or structural characteristic of TRICH.

The invention also encompasses polynucleotides which encode TRICH. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:33-64, which encodes TRICH. The polynucleotide sequences of SEQ ID NO:33-64, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding TRICH. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding TRICH. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:33-64 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:33-64. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of TRICH.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding TRICH, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring TRICH, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode TRICH and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring TRICH under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding TRICH or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide

occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding TRICH and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode TRICH and TRICH derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding TRICH or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:33-64 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (Applied Biosystems), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (Applied Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (Applied Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding TRICH may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic

DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 primer analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, Applied Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode TRICH may be cloned in recombinant DNA molecules that direct expression of TRICH, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of

the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express TRICH.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter TRICH-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of TRICH, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding TRICH may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, TRICH itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY; pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (Applied Biosystems). Additionally, the amino acid sequence of TRICH, or any part thereof, may be altered during direct synthesis and/or combined with sequences

from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active TRICH, the nucleotide sequences encoding TRICH or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding TRICH. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding TRICH. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding TRICH and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding TRICH and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding TRICH. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or

animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al. (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding TRICH. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding TRICH can be achieved using a multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding TRICH into the vector's multiple cloning site disrupts the *lac*Z gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of TRICH are needed, e.g. for the production of antibodies, vectors which direct high level expression of TRICH may be used. For example, vectors containing the strong, inducible SP6 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of TRICH. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; and Scorer, C.A. et al. (1994) Bio/Technology 12:181-184.)

Plant systems may also be used for expression of TRICH. Transcription of sequences encoding TRICH may be driven by viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311).

Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; and Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding TRICH may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses TRICH in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of TRICH in cell lines is preferred. For example, sequences encoding TRICH can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*$^-$ and *apr*$^-$ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980)

Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding TRICH is inserted within a marker gene sequence, transformed cells containing sequences encoding TRICH can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding TRICH under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding TRICH and that express TRICH may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of TRICH using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on TRICH is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding TRICH include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding TRICH, or any fragments thereof, may be cloned into a vector for

the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding TRICH may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode TRICH may be designed to contain signal sequences which direct secretion of TRICH through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding TRICH may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric TRICH protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of TRICH activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize

these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the TRICH encoding sequence and the heterologous protein sequence, so that TRICH may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled TRICH may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

TRICH of the present invention or fragments thereof may be used to screen for compounds that specifically bind to TRICH. At least one and up to a plurality of test compounds may be screened for specific binding to TRICH. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of TRICH, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, e.g., Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which TRICH binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express TRICH, either as a secreted protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing TRICH or cell membrane fractions which contain TRICH are then contacted with a test compound and binding, stimulation, or inhibition of activity of either TRICH or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with TRICH, either in solution or affixed to a solid support, and detecting the binding of TRICH to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

TRICH of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of TRICH. Such compounds may include agonists, antagonists, or partial

or inverse agonists. In one embodiment, an assay is performed under conditions permissive for TRICH activity, wherein TRICH is combined with at least one test compound, and the activity of TRICH in the presence of a test compound is compared with the activity of TRICH in the absence of the test compound. A change in the activity of TRICH in the presence of the test compound is indicative of a compound that modulates the activity of TRICH. Alternatively, a test compound is combined with an in vitro or cell-free system comprising TRICH under conditions suitable for TRICH activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of TRICH may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding TRICH or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent Number 5,175,383 and U.S. Patent Number 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding TRICH may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding TRICH can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding TRICH is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with

potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress TRICH, e.g., by secreting TRICH in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of TRICH and transporters and ion channels. In addition, the expression of TRICH is closely associated with adrenal, testicular, and prostate tumors, Crohn's disease, teratocarcinoma and dendritic cells, brain, lung, ileum, small intestine, uterine myometrial, colon, and pancreatic tissues. Therefore, TRICH appears to play a role in transport, neurological, muscle, immunological, and cell proliferative disorders. In the treatment of disorders associated with increased TRICH expression or activity, it is desirable to decrease the expression or activity of TRICH. In the treatment of disorders associated with decreased TRICH expression or activity, it is desirable to increase the expression or activity of TRICH.

Therefore, in one embodiment, TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup disease, and Fanconi disease; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other

extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromusoular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, and acid maltase deficiency (AMD, also known as Pompe's disease); an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal

circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing TRICH or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those described above.

In a further embodiment, a composition comprising a substantially purified TRICH in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of TRICH may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of TRICH including, but not limited to, those listed above.

In a further embodiment, an antagonist of TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH. Examples of such disorders include, but are not limited to, those transport, neurological, muscle, immunological, and cell proliferative disorders described above. In one aspect, an antibody which specifically binds TRICH may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express TRICH.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding TRICH may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of TRICH including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of TRICH may be produced using methods which are generally known in the art. In particular, purified TRICH may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind TRICH. Antibodies to TRICH may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with TRICH or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to TRICH have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of TRICH amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to TRICH may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce TRICH-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be

generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for TRICH may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or mohoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between TRICH and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering TRICH epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for TRICH. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of TRICH-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple TRICH epitopes, represents the average affinity, or avidity, of the antibodies for TRICH. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular TRICH epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the TRICH-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of TRICH, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of TRICH-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al. supra.)

In another embodiment of the invention, the polynucleotides encoding TRICH, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding TRICH. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding TRICH. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding TRICH may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial

hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and N. Somia (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in TRICH expression or regulation causes disease, the expression of TRICH from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in TRICH are treated by constructing mammalian expression vectors encoding TRICH and introducing these vectors by mechanical means into TRICH-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of TRICH include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). TRICH may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V. and Blau, H.M. supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding TRICH from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver

polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to TRICH expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding TRICH under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding TRICH to cells which have one or more genetic abnormalities with respect to the expression of TRICH. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu.

Rev. Nutr. 19:511-544 and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding TRICH to target cells which have one or more genetic abnormalities with respect to the expression of TRICH. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing TRICH to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res. 169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding TRICH to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotechnol. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for TRICH into the alphavirus genome in place of the capsid-coding region results in the production of a large number of TRICH-coding RNAs and the synthesis of high levels of TRICH in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic replication of alphaviruses can be altered to suit the needs of the gene therapy application

(Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of TRICH into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding TRICH.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding TRICH. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding TRICH. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased TRICH expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding TRICH may be therapeutically useful, and in the treatment of disorders associated with decreased TRICH expression or activity, a compound which specifically promotes expression of the polynucleotide encoding TRICH may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding TRICH is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding TRICH are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding TRICH. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of

the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such compositions may consist of TRICH, antibodies to TRICH, and mimetics, agonists, antagonists, or inhibitors of TRICH.

The compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g., Patton, J.S. et al., U.S.

Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of compositions may be prepared for direct intracellular delivery of macromolecules comprising TRICH or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, TRICH or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example TRICH or fragments thereof, antibodies of TRICH, and agonists, antagonists or inhibitors of TRICH, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy.

Long-acting compositions may be administered every 3 to 4 days, every week, or biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind TRICH may be used for the diagnosis of disorders characterized by expression of TRICH, or in assays to monitor patients being treated with TRICH or agonists, antagonists, or inhibitors of TRICH. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for TRICH include methods which utilize the antibody and a label to detect TRICH in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring TRICH, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of TRICH expression. Normal or standard values for TRICH expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibodies to TRICH under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of TRICH expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding TRICH may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of TRICH may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of TRICH, and to monitor regulation of TRICH levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding TRICH or closely related molecules may be used to identify nucleic acid sequences which encode TRICH. The specificity of the probe, whether it is made

from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the probe identifies only naturally occurring sequences encoding TRICH, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the TRICH encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:33-64 or from genomic sequences including promoters, enhancers, and introns of the TRICH gene.

Means for producing specific hybridization probes for DNAs encoding TRICH include the cloning of polynucleotide sequences encoding TRICH or TRICH derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding TRICH may be used for the diagnosis of disorders associated with expression of TRICH. Examples of such disorders include, but are not limited to, a transport disorder such as akinesia, amyotrophic lateral sclerosis, ataxia telangiectasia, cystic fibrosis, Becker's muscular dystrophy, Bell's palsy, Charcot-Marie Tooth disease, diabetes mellitus, diabetes insipidus, diabetic neuropathy, Duchenne muscular dystrophy, hyperkalemic periodic paralysis, normokalemic periodic paralysis, Parkinson's disease, malignant hyperthermia, multidrug resistance, myasthenia gravis, myotonic dystrophy, catatonia, tardive dyskinesia, dystonias, peripheral neuropathy, cerebral neoplasms, prostate cancer, cardiac disorders associated with transport, e.g., angina, bradyarrythmia, tachyarrythmia, hypertension, Long QT syndrome, myocarditis, cardiomyopathy, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, thyrotoxic myopathy, ethanol myopathy, dermatomyositis, inclusion body myositis, infectious myositis, polymyositis, neurological disorders associated with transport, e.g., Alzheimer's disease, amnesia, bipolar disorder, dementia, depression, epilepsy, Tourette's disorder, paranoid psychoses, and schizophrenia, and other disorders associated with transport, e.g., neurofibromatosis, postherpetic neuralgia, trigeminal neuropathy, sarcoidosis, sickle cell anemia, Wilson's disease, cataracts, infertility, pulmonary artery stenosis, sensorineural autosomal deafness, hyperglycemia, hypoglycemia, Grave's disease, goiter, Cushing's disease, Addison's disease, glucose-galactose malabsorption syndrome, hypercholesterolemia, adrenoleukodystrophy, Zellweger syndrome, Menkes disease, occipital horn syndrome, von Gierke disease, cystinuria, iminoglycinuria, Hartup disease, and

Fanconi disease; a neurological disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system including Down syndrome, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, seasonal affective disorder (SAD), akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, Tourette's disorder, progressive supranuclear palsy, corticobasal degeneration, and familial frontotemporal dementia; a muscle disorder such as cardiomyopathy, myocarditis, Duchenne's muscular dystrophy, Becker's muscular dystrophy, myotonic dystrophy, central core disease, nemaline myopathy, centronuclear myopathy, lipid myopathy, mitochondrial myopathy, infectious myositis, polymyositis, dermatomyositis, inclusion body myositis, thyrotoxic myopathy, ethanol myopathy, angina, anaphylactic shock, arrhythmias, asthma, cardiovascular shock, Cushing's syndrome, hypertension, hypoglycemia, myocardial infarction, migraine, pheochromocytoma, and myopathies including encephalopathy, epilepsy, Kearns-Sayre syndrome, lactic acidosis, myoclonic disorder, ophthalmoplegia, and acid maltase deficiency (AMD, also known as Pompe's disease); an immunological disorder such as acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, autoimmune polyendocrinopathy-candidiasis-ectodermal dystrophy (APECED), bronchitis, cholecystitis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, episodic lymphopenia with lymphocytotoxins, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, hypereosinophilia, irritable bowel syndrome, multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, osteoarthritis, osteoporosis, pancreatitis, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic

anaphylaxis, systemic lupus erythematosus, systemic sclerosis, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, viral, bacterial, fungal, parasitic, protozoal, and helminthic infections, and trauma; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding TRICH may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered TRICH expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding TRICH may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding TRICH may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding TRICH in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of TRICH, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding TRICH, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding TRICH may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding TRICH, or a fragment of a polynucleotide complementary to the polynucleotide encoding TRICH, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding TRICH may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding TRICH are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the

alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of TRICH include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described below. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, TRICH, fragments of TRICH, or antibodies specific for TRICH may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity (Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are

analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for TRICH to quantify the levels of TRICH expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference

in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g., Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding TRICH may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, for example, Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding TRICH on a physical

map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, TRICH, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between TRICH and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with TRICH, or fragments thereof, and washed. Bound TRICH is then detected by methods well known in the art. Purified TRICH can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding TRICH specifically compete with a test compound for binding TRICH. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with TRICH.

In additional embodiments, the nucleotide sequences which encode TRICH may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below including U.S. Ser. No. 60/216,547, U.S. Ser. No. 60/218,232, U.S. Ser. No. 60/220,112, and U.S. Ser. No. 60/221,839 are expressly incorporated by reference herein., are expressly incorporated by reference herein.

EXAMPLES

I. Construction of cDNA Libraries

Incyte cDNAs were derived from cDNA libraries described in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA) and shown in Table 4, column 5. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A)+ RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g.,

PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), PCDNA2.1 plasmid (Invitrogen, Carlsbad CA), PBK-CMV plasmid (Stratagene), or pINCY (Incyte Genomics, Palo Alto CA), or derivatives thereof. Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (Applied Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (Applied Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VIII.

The polynucleotide sequences derived from Incyte cDNAs were validated by removing vector, linker, and poly(A) sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programming, and dinucleotide nearest neighbor analysis. The Incyte cDNA sequences or translations thereof were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and hidden Markov model (HMM)-based protein family databases such as PFAM. (HMM is a probabilistic approach which analyzes consensus primary structures of gene families. See, for example, Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.) The queries were performed using programs based on BLAST, FASTA, BLIMPS, and HMMER. The Incyte cDNA sequences were assembled to produce full length polynucleotide sequences. Alternatively, GenBank cDNAs, GenBank ESTs, stitched sequences, stretched sequences, or Genscan-predicted coding sequences (see Examples IV and V) were used to extend Incyte cDNA assemblages to full length. Assembly was performed using programs based on Phred, Phrap, and Consed, and cDNA assemblages were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length polypeptide sequences. Alternatively, a polypeptide of the invention may begin at any of the methionine residues of the full length translated polypeptide. Full length polypeptide sequences were subsequently analyzed by querying against databases such as the GenBank protein databases (genpept), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and hidden Markov model (HMM)-based protein family databases such as PFAM. Full length polynucleotide sequences are also analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments are generated using default parameters specified by the CLUSTAL algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

Table 7 summarizes the tools, programs, and algorithms used for the analysis and assembly of Incyte cDNA and full length sequences and provides applicable descriptions, references, and threshold parameters. The first column of Table 7 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score or the lower the probability value, the greater the identity between two sequences).

The programs described above for the assembly and analysis of full length polynucleotide and polypeptide sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:33-64. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies are described in Table 4, column 4.

IV. Identification and Editing of Coding Sequences from Genomic DNA

Putative transporters and ion channels were initially identified by running the Genscan gene identification program against public genomic sequence databases (e.g., gbpri and gbhtg). Genscan is a general-purpose gene identification program which analyzes genomic DNA sequences from a variety of organisms (See Burge, C. and S. Karlin (1997) J. Mol. Biol. 268:78-94, and Burge, C. and S. Karlin (1998) Curr. Opin. Struct. Biol. 8:346-354). The program concatenates predicted exons to form an assembled cDNA sequence extending from a methionine to a stop codon. The output of Genscan is a FASTA database of polynucleotide and polypeptide sequences. The maximum range of sequence for Genscan to analyze at once was set to 30 kb. To determine which of these Genscan predicted cDNA sequences encode transporters and ion channels, the encoded polypeptides were analyzed by querying against PFAM models for transporters and ion channels. Potential transporters and ion channels were also identified by homology to Incyte cDNA sequences that had been annotated as transporters and ion channels. These selected Genscan-predicted sequences were then compared by BLAST analysis to the genpept and gbpri public databases. Where necessary, the Genscan-predicted sequences were then edited by comparison to the top BLAST hit from genpept to correct errors in the sequence predicted by Genscan, such as extra or omitted exons. BLAST analysis was also used to find any Incyte cDNA or public cDNA coverage of the Genscan-predicted sequences, thus providing evidence for transcription. When Incyte cDNA coverage was available, this information was used to correct or confirm the Genscan predicted sequence. Full length polynucleotide sequences were obtained by assembling Genscan-predicted coding sequences with Incyte cDNA sequences and/or public cDNA sequences using the assembly process described in Example III. Alternatively, full length polynucleotide sequences were derived entirely from edited or unedited Genscan-predicted coding sequences.

V. Assembly of Genomic Sequence Data with cDNA Sequence Data

"Stitched" Sequences

Partial cDNA sequences were extended with exons predicted by the Genscan gene identification program described in Example IV. Partial cDNAs assembled as described in Example III were mapped to genomic DNA and parsed into clusters containing related cDNAs and Genscan exon predictions from one or more genomic sequences. Each cluster was analyzed using an algorithm based on graph theory and dynamic programming to integrate cDNA and genomic information, generating possible splice variants that were subsequently confirmed, edited, or extended to create a full length sequence.

Sequence intervals in which the entire length of the interval was present on more than one sequence in the cluster were identified, and intervals thus identified were considered to be equivalent by transitivity. For example, if an interval was present on a cDNA and two genomic sequences, then all three intervals were considered to be equivalent. This process allows unrelated but consecutive genomic sequences to be brought together, bridged by cDNA sequence. Intervals thus identified were then "stitched" together by the stitching algorithm in the order that they appear along their parent sequences to generate the longest possible sequence, as well as sequence variants. Linkages between intervals which proceed along one type of parent sequence (cDNA to cDNA or genomic sequence to genomic sequence) were given preference over linkages which change parent type (cDNA to genomic sequence). The resultant stitched sequences were translated and compared by BLAST analysis to the genpept and gbpri public databases. Incorrect exons predicted by Genscan were corrected by comparison to the top BLAST hit from genpept. Sequences were further extended with additional cDNA sequences, or by inspection of genomic DNA, when necessary.

"Stretched" Sequences

Partial DNA sequences were extended to full length with an algorithm based on BLAST analysis. First, partial cDNAs assembled as described in Example III were queried against public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases using the BLAST program. The nearest GenBank protein homolog was then compared by BLAST analysis to either Incyte cDNA sequences or GenScan exon predicted sequences described in Example IV. A chimeric protein was generated by using the resultant high-scoring segment pairs (HSPs) to map the translated sequences onto the GenBank protein homolog. Insertions or deletions may occur in the chimeric protein with respect to the original GenBank protein homolog. The GenBank protein homolog, the chimeric protein, or both were used as probes to search for homologous genomic sequences from the public human genome databases. Partial DNA sequences were therefore "stretched" or extended by the addition of homologous genomic sequences. The resultant stretched sequences were examined to determine whether it contained a complete gene.

VI. Chromosomal Mapping of TRICH Encoding Polynucleotides

The sequences which were used to assemble SEQ ID NO:33-64 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:33-64 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 7). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences

had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

Map locations are represented by ranges, or intervals, of human chromosomes. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters. Human genome maps and other resources available to the public, such as the NCBI "GeneMap'99" World Wide Web site (http://www.ncbi.nlm.nih.gov/genemap/), can be employed to determine if previously identified disease genes map within or in proximity to the intervals indicated above.

VII. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, supra, ch. 7; Ausubel (1995) supra, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score} \times \text{Percent Identity}}{5 \times \text{minimum}\{\text{length(Seq. 1)}, \text{length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For

example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and 70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

Alternatively, polynucleotide sequences encoding TRICH are analyzed with respect to the tissue sources from which they were derived. For example, some full length sequences are assembled, at least in part, with overlapping Incyte cDNA sequences (see Example III). Each cDNA sequence is derived from a cDNA library constructed from a human tissue. Each human tissue is classified into one of the following organ/tissue categories: cardiovascular system; connective tissue; digestive system; embryonic structures; endocrine system; exocrine glands; genitalia, female; genitalia, male; germ cells; hemic and immune system; liver; musculoskeletal system; nervous system; pancreas; respiratory system; sense organs; skin; stomatognathic system; unclassified/mixed; or urinary tract. The number of libraries in each category is counted and divided by the total number of libraries across all categories. Similarly, each human tissue is classified into one of the following disease/condition categories: cancer, cell line, developmental, inflammation, neurological, trauma, cardiovascular, pooled, and other, and the number of libraries in each category is counted and divided by the total number of libraries across all categories. The resulting percentages reflect the tissue- and disease-specific expression of cDNA encoding TRICH. cDNA sequences and cDNA library/tissue information are found in the LIFESEQ GOLD database (Incyte Genomics, Palo Alto CA).

VIII. Extension of TRICH Encoding Polynucleotides

Full length polynucleotide sequences were also produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer was synthesized to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and 2-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme

(Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 μl to 10 μl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (Applied Biosystems).

In like manner, full length polynucleotide sequences are verified using the above procedure or are used to obtain 5' regulatory sequences using the above procedure along with oligonucleotides designed for such extension, and an appropriate genomic library.

IX. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:33-64 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

X. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra.), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/µl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37° C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85° C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 µg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 µl of the array element DNA, at an average concentration of 100 ng/µl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60° C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 µl of sample mixture consisting of 0.2 µg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65° C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 µl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60° C. The arrays are washed for 10 min at 45° C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45° C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate

filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

XI. Complementary Polynucleotides

Sequences complementary to the TRICH-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring TRICH. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of TRICH. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the TRICH-encoding transcript.

XII. Expression of TRICH

Expression and purification of TRICH is achieved using bacterial or virus-based expression systems. For expression of TRICH in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express TRICH upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of TRICH in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding TRICH by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, TRICH is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from TRICH at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified TRICH obtained by these methods can be used directly in the assays shown in Examples XVI, XVII, and XVIII where applicable.

XIII. Functional Assays

TRICH function is assessed by expressing the sequences encoding TRICH at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include PCMV SPORT (Life Technologies) and PCR3.1 (Invitrogen, Carlsbad CA), both of which

contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of TRICH on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding TRICH and either CD64 or CD64-GFP. CD64 and CD64-GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding TRICH and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIV. Production of TRICH Specific Antibodies

TRICH substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the TRICH amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (Applied Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-TRICH activity by, for example, binding the peptide or TRICH to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XV. Purification of Naturally Occurring TRICH Using Specific Antibodies

Naturally occurring or recombinant TRICH is substantially purified by immunoaffinity chromatography using antibodies specific for TRICH. An immunoaffinity column is constructed by covalently coupling anti-TRICH antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing TRICH are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of TRICH (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/TRICH binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and TRICH is collected.

XVI. Identification of Molecules Which Interact with TRICH

Molecules which interact with TRICH may include transporter substrates, agonists or antagonists, modulatory proteins such as Gβγ proteins (Reimann, supra) or proteins involved in TRICH localization or clustering such as MAGUKs (Craven, supra). TRICH, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent. (See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled TRICH, washed, and any wells with labeled TRICH complex are assayed. Data obtained using different concentrations of TRICH are used to calculate values for the number, affinity, and association of TRICH with the candidate molecules.

Alternatively, proteins that interact with TRICH are isolated using the yeast 2-hybrid system (Fields, S. and O. Song (1989) Nature 340:245-246). TRICH, or fragments thereof, are expressed as fusion proteins with the DNA binding domain of Gal4 or lexA, and potential interacting proteins are expressed as fusion proteins with an activation domain. Interactions between the TRICH fusion protein and the TRICH interacting proteins (fusion proteins with an activation domain) reconstitute a transactivation function that is observed by expression of a reporter gene. Yeast 2-hybrid systems are

commercially available, and methods for use of the yeast 2-hybrid system with ion channel proteins are discussed in Niethammer, M. and M. Sheng (1998, Meth. Enzymol. 293:104-122).

TRICH may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Potential TRICH agonists or antagonists may be tested for activation or inhibition of TRICH ion channel activity using the assays described in section XVIII.

XVII. Demonstration of TRICH Activity

Ion channel activity of TRICH is demonstrated using an electrophysiological assay for ion conductance. TRICH can be expressed by transforming a mammalian cell line such as COS7, HeLa or CHO with a eukaryotic expression vector encoding TRICH. Eukaryotic expression vectors are commercially available, and the techniques to introduce them into cells are well known to those skilled in the art. A second plasmid which expresses any one of a number of marker genes, such as ß-galactosidase, is co-transformed into the cells to allow rapid identification of those cells which have taken up and expressed the foreign DNA. The cells are incubated for 48-72 hours after transformation under conditions appropriate for the cell line to allow expression and accumulation of TRICH and ß-galactosidase.

Transformed cells expressing ß-galactosidase are stained blue when a suitable colorimetric substrate is added to the culture media under conditions that are well known in the art. Stained cells are tested for differences in membrane conductance by electrophysiological techniques that are well known in the art. Untransformed cells, and/or cells transformed with either vector sequences alone or ß-galactosidase sequences alone, are used as controls and tested in parallel. Cells expressing TRICH will have higher anion or cation conductance relative to control cells. The contribution of TRICH to conductance can be confirmed by incubating the cells using antibodies specific for TRICH. The antibodies will bind to the extracellular side of TRICH, thereby blocking the pore in the ion channel, and the associated conductance.

Alternatively, ion channel activity of TRICH is measured as current flow across a TRICH-containing Xenopus laevis oocyte membrane using the two-electrode voltage-clamp technique (Ishi et al., supra; Jegla, T. and L. Salkoff (1997) J. Neurosci. 17:32-44). TRICH is subcloned into an appropriate Xenopus oocyte expression vector, such as pBF, and 0.5-5 ng of mRNA is injected into mature stage IV oocytes. Injected oocytes are incubated at 18°C for 1-5 days. Inside-out macropatches are excised into an intracellular solution containing 116 mM K-gluconate, 4 mM KCl, and 10 mM Hepes (pH 7.2). The intracellular solution is supplemented with varying concentrations of the TRICH mediator, such as cAMP, cGMP, or Ca^{+2} (in the form of $CaCl_2$), where appropriate.

Electrode resistance is set at 2-5 MΩ and electrodes are filled with the intracellular solution lacking mediator. Experiments are performed at room temperature from a holding potential of 0 mV. Voltage ramps (2.5 s) from -100 to 100 mV are acquired at a sampling frequency of 500 Hz. Current measured is proportional to the activity of TRICH in the assay.

In particular, the activities of TRICH-1, TRICH-2, and TRICH-10, are measured as K^+ conductance, the activities of TRICH-6 and TRICH-9 are measured as K^+ conductance in the presence of membrane stretch or free fatty acids, the activities of TRICH-18, TRICH-25 and TRICH-31 are measured as voltage-gated K^+ conductance, TRICH-5 activity is measured as Cl^- conductance in the presence of GABA, TRICH-11 activity is measured as cation conductance in the presence of heat, and the activity of TRICH-9, TRICH-28 is measured as Ca^{2+} conductance.

Transport activity of TRICH is assayed by measuring uptake of labeled substrates into Xenopus laevis oocytes. Oocytes at stages V and VI are injected with TRICH mRNA (10 ng per oocyte) and incubated for 3 days at 18°C in OR2 medium (82.5mM NaCl, 2.5 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 1mM Na_2HPO_4, 5 mM Hepes, 3.8 mM NaOH , 50μg/ml gentamycin, pH 7.8) to allow expression of TRICH. Oocytes are then transferred to standard uptake medium (100mM NaCl, 2 mM KCl, 1mM $CaCl_2$, 1mM $MgCl_2$, 10 mM Hepes/Tris pH 7.5). Uptake of various substrates (e.g., amino acids, sugars, drugs, ions, and neurotransmitters) is initiated by adding labeled substrate (e.g. radiolabeled with ^3H, fluorescently labeled with rhodamine, etc.) to the oocytes. After incubating for 30 minutes, uptake is terminated by washing the oocytes three times in Na^+-free medium, measuring the incorporated label, and comparing with controls. TRICH activity is proportional to the level of internalized labeled substrate. In particular, test substrates include pigment precursors and related molecules for TRICH-3, aminophospholipids for TRICH-4, fructose and glucose for TRICH-7 and TRICH-15, amino acids for TRICH-8, Na^+ and iodide for TRICH-12, Na^+ and H^+ for TRICH-13 and TRICH-21, Na^+ and glucose for TRICH-16 and TRICH-19, and glucose for TRICH-23, TRICH-26, TRICH-29, TRICH-30, and TRICH-32.

ATPase activity associated with TRICH can be measured by hydrolysis of radiolabeled ATP-[γ-^{32}P], separation of the hydrolysis products by chromatographic methods, and quantitation of the recovered ^{32}P using a scintillation counter. The reaction mixture contains ATP-[γ-^{32}P] and varying amounts of TRICH in a suitable buffer incubated at 37°C for a suitable period of time. The reaction is terminated by acid precipitation with trichloroacetic acid and then neutralized with base, and an aliquot of the reaction mixture is subjected to membrane or filter paper-based chromatography to separate the reaction products. The amount of ^{32}P liberated is counted in a scintillation counter. The amount of radioactivity recovered is proportional to the ATPase activity of TRICH in the assay.

XVIII. Identification of TRICH Agonists and Antagonists

TRICH is expressed in a eukaryotic cell line such as CHO (Chinese Hamster Ovary) or HEK (Human Embryonic Kidney) 293. Ion channel activity of the transformed cells is measured in the presence and absence of candidate agonists or antagonists. Ion channel activity is assayed using patch clamp methods well known in the art or as described in Example XVII. Alternatively, ion channel activity is assayed using fluorescent techniques that measure ion flux across the cell membrane (Velicelebi, G. et al. (1999) Meth. Enzymol. 294:20-47; West, M.R. and C.R. Molloy (1996) Anal. Biochem. 241:51-58). These assays may be adapted for high-throughput screening using microplates. Changes in internal ion concentration are measured using fluorescent dyes such as the Ca^{2+} indicator Fluo-4 AM, sodium-sensitive dyes such as SBFI and sodium green, or the Cl^- indicator MQAE (all available from Molecular Probes) in combination with the FLIPR fluorimetric plate reading system (Molecular Devices). In a more generic version of this assay, changes in membrane potential caused by ionic flux across the plasma membrane are measured using oxonyl dyes such as $DiBAC_4$ (Molecular Probes). $DiBAC_4$ equilibrates between the extracellular solution and cellular sites according to the cellular membrane potential. The dye's fluorescence intensity is 20-fold greater when bound to hydrophobic intracellular sites, allowing detection of $DiBAC_4$ entry into the cell (Gonzalez, J.E. and P.A. Negulescu (1998) Curr. Opin. Biotechnol. 9:624-631). Candidate agonists or antagonists may be selected from known ion channel agonists or antagonists, peptide libraries, or combinatorial chemical libraries.

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Incyte Project ID	Polypeptide SEQ ID NO:	Incyte Polypeptide ID	Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID
3474673	1	3474673CD1	33	3474673CB1
4588877	2	4588877CD1	34	4588877CB1
7472214	3	7472214CD1	35	7472214CB1
7473053	4	7473053CD1	36	7473053CB1
7473347	5	7473347CD1	37	7473347CB1
7474240	6	7474240CD1	38	7474240CB1
7475338	7	7475338CD1	39	7475338CB1
7476747	8	7476747CD1	40	7476747CB1
7477898	9	7477898CD1	41	7477898CB1
7472728	10	7472728CD1	42	7472728CB1
7474322	11	7474322CD1	43	7474322CB1
5455621	12	5455621CD1	44	5455621CB1
7477248	13	7477248CD1	45	7477248CB1
2944004	14	2944004CD1	46	2944004CB1
3046849	15	3046849CD1	47	3046849CB1
4538363	16	4538363CD1	48	4538363CB1
6427460	17	6427460CD1	49	6427460CB1
7474127	18	7474127CD1	50	7474127CB1
7476949	19	7476949CD1	51	7476949CB1
7477249	20	7477249CD1	52	7477249CB1
7477720	21	7477720CD1	53	7477720CB1
7477852	22	7477852CD1	54	7477852CB1
1471717	23	1471717CD1	55	1471717CB1
3874406	24	3874406CD1	56	3874406CB1
4599654	25	4599654CD1	57	4599654CB1
5047435	26	5047435CD1	58	5047435CB1
7475603	27	7475603CD1	59	7475603CB1
7477845	28	7477845CD1	60	7477845CB1
168827	29	168827CD1	61	168827CB1
7472734	30	7472734CD1	62	7472734CB1
7473473	31	7473473CD1	63	7473473CB1
7477725	32	7477725CD1	64	7477725CB1

Table 2

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
1	3474673CD1	g13507377	1.00E-151	[fl][Homo sapiens] potassium channel TASK-4 (Decher,N. et al. (2001) FEBS Lett. 492 (1-2), 84-89)
2	4588877CD1	g13926111	3.00E-96	[fl][Homo sapiens] (AF358910) 2P domain potassium channel Talk-2
3	7472214CD1	g1107730	1.70E-243	[Mus musculus] ABC8 (Savary,S. et al. (1996) Mamm. Genome 7 (9), 673-676)
		g11342541	0	[fl][Homo sapiens] putative white family ATP-binding cassette transporter
4	7473053CD1	g3850108	9.00E-209	[Schizosaccharomyces pombe] putative calcium-transporting atpase
		g3628757	0	[Homo sapiens] FIC1 (Bull,L.N. et al. (1998) Nat. Genet. 18 (3), 219-224)
5	7473347CD1	g1060975	1.70E-206	[Rattus norvegicus] GABA receptor rho-3 subunit precursor (Ogurusu,T. et al. (1996) Biochim. Biophys. Acta 1305 (1-2), 15-18)
6	7474240CD1	g2745727	0	[Rattus norvegicus] potassium channel (Shi,W. et al. (1997) J. Neurosci. 17 (24), 9423-9432)
7	7475338CD1	g183298	2.10E-158	[Homo sapiens] GLUT5 protein (Kayano,T. et al. (1990) J. Biol. Chem. 265 (22), 13276-13282)
9	7477898CD1	g2745729	0	[Rattus norvegicus] potassium channel (Shi,W. et al. (1997) J. Neurosci. 17 (24), 9423-9432)
10	7472728CD1	g8452900	3.50E-261	[Rattus norvegicus] potassium channel TREK-2 (Bang,H. et al. (2000) J. Biol. Chem. 275 (23), 17412-17419)
11	7474322CD1	g12003146	0	[fl][Homo sapiens] capsaicin receptor
12	5455621CD1	g1399954	3.00E-143	[Rattus norvegicus] thyroid sodium/iodide symporter NIS (Dai,G. et al. (1996) Nature 379 (6564), 458-460)
13	7477248CD1	g2944233	3.10E-195	[Homo sapiens] sodium-hydrogen exchanger 6 (Numata,M. et al. (1998) J. Biol. Chem. 273 (12), 6951-6959)
14	2944004CD1	g3451312	1.40E-188	[Schizosaccharomyces pombe] membrane atpase
15	3046849CD1	g12802047	0	[fl][Homo sapiens] (AJ271290) facilitative glucose transporter GLUT11

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
16	4538363CD1	g338055	7.40E-181	[Homo sapiens] Na+/glucose cotransporter (Hediger,M.A. et al. (1989) Proc. Natl. Acad. Sci. U.S.A. 86 (15), 5748-5752)
17	6427460CD1	g6457274	0	[Mus musculus] putative E1-E2 ATPase (Halleck,M.S. et al. (1999) Physiol. Genomics (Online) 1 (3), 139-150)
18	7474127CD1	g206044	0	[Rattus norvegicus] potassium channel Kv3.2b (Wiedmann,R. et al. (1991) FEBS Lett. 288, 163-167)
19	7476949CD1	g9588428	0	[5' incom][Homo sapiens] dJ1024N4.1 (novel Sodium:solute symporter family member similar to SLC5A1 (SGLT1))
		g338055	3.70E-202	[Homo sapiens] Na+/glucose cotransporter (Hediger,M.A. et al. (1989) Proc. Natl. Acad. Sci. U.S.A. 86 (15), 5748-5752)
20	7477249CD1	g7715417	0	[Oryctolagus cuniculus] RING-finger binding protein (Mansharamani,M. et al. (2001) J. Biol. Chem. 276 (5), 3641-3649)
21	7477720CD1	g205709	0	[Rattus norvegicus] sodium-hydrogen exchange protein-isoform 4 (Orlowski,J. et al. (1992) J. Biol. Chem. 267, 9331-9339)
22	7477852CD1	g8920219	0	[fl][Homo sapiens] epithelial calcium channel (Muller,D. et al. (2000) Genomics 67 (1), 48-53)
23	1471717CD1	g529590	5.00E-36	[Rattus norvegicus] liver-specific transport protein (Simonson,G.D. et al. (1994) J. Cell. Sci 107, 1065-1072)
24	3874406CD1	g1514530	1.90E-117	[Homo sapiens] ABC-C transporter (Klugbauer,N. et al. (1996) FEBS Lett. 391 (1-2), 61-65)
25	4599654CD1	g3242244	0	[Mus musculus] hyperpolarization-activated cation channel, HAC3 (Ludwig,A. et al. (1998) Nature 393 (6685), 587-591)

Table 2 (cont.)

Polypeptide SEQ ID NO:	Incyte Polypeptide ID	GenBank ID NO:	Probability score	GenBank Homolog
26	5047435CD1	g13445575	0	[fl][Homo sapiens] facilitative glucose transporter GLUT10 (McVie-Wylie,A.J. et al. (2001) Genomics 72 (1), 113-117)
27	7475603CD1	g9211112	0	[fl][Homo sapiens] macrophage ABC transporter (Kaminski,W.E. et al. (2000) Biochem. Biophys. Res. Commun. 273 (2), 532-538)
28	7477845CD1	g3800830	0	[Rattus norvegicus] putative four repeat ion channel (Lee,J.H. et al. (1999) FEBS Lett. 445 (2-3), 231-236)
29	168827CD1	g7707622	1.20E-116	[Homo sapiens] organic anion transporter 4 (Cha,S.H. et al. (2000) J. Biol. Chem. 275 (6), 4507-4512)
		g3004482	0	[fl][Rattus norvegicus] putative integral membrane transport protein (Schomig,E. et al. (1998) FEBS Lett. 425 (1), 79-86)
30	7472734CD1	g7707622	4.50E-117	[Homo sapiens] organic anion transporter 4 (Cha,S.H. et al. (2000) J. Biol. Chem. 275 (6), 4507-4512)
		g3004482	0	[fl][Rattus norvegicus] putative integral membrane transport protein (Schomig,E. et al. (1998) FEBS Lett. 425 (1), 79-86)
31	7473473CD1	g6625694	0	[Rattus norvegicus] potasium channel Eag2 (Saganich,M.J. et al. (1999) J. Neurosci. 19 (24), 10789-10802)
32	7477725CD1	g3004482	1.00E-177	[fl][Rattus norvegicus] putative integral membrane transport protein (Schomig,E. et al. (1998) FEBS Lett. 425 (1), 79-86)
		g7707622	4.20E-130	[Homo sapiens] organic anion transporter 4 (Cha,S.H. et al. (2000) J. Biol. Chem. 275 (6), 4507-4512)

Table 3

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
1	3474673CD1	332	S201 S207 S234 S265 S280 S281 S289 S51 T169 T67	N65 N94	Transmembrane domains: R130-M155, V245-L264	HMMER
					TASK K+ channel domain: V14-S332	HMMER_PFAM
2	4588877CD1	226	S101 S128 S159 S174 S175 S183 S95		Transmembrane domain: V139-L158	HMMER
					CHANNEL PROTEIN IONIC POTASSIUM SUBUNIT K+ PUTATIVE SUBFAMILY K MEMBER PD021430:-A78-E162	BLAST_PRODOM
3	7472214CD1	646	S143 S229 S261 S340 S341 S463 S554 S57 S644 S69 S89 T138 T157 T23 T472 T500 T591	N169 N422	Transmembrane domains: S430-M450, W564-D589, M618-V637	HMMER
					ABC transporter domain: R95-G277	HMMER_PFAM
					ABC transporters family signature BL00211: I100-F111, L201-D232	BLIMPS_BLOCKS
					ABC transporters family signature: V181-D232	PROFILESCAN
					PROTEIN TRANSMEMBRANE TRANSPORT ATPBINDING TRANSPORTER MEMBRANE ABC GLYCOPROTEIN INNER PUTATIVE PD000633: T365-Y583	BLAST_PRODOM
					do WHITE; FRUIT; FLY; SCARLET; DM05200|P45844|289-650: G277-L623	BLAST_DOMO
					ABC TRANSPORTERS FAMILY DM00008|P45844|73-287: I61-Q276	BLAST_DOMO
					ABC transporter motif: L201-L215	MOTIFS
					ATP/GTP binding site (P-loop): G102-S109	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
4	7473053CD1	1190	S153 S259 S268 S391 S413 S452 S493 S545 S573 S624 S631 S687 S723 S739 S744 S832 S1174 S1132 S1164 S1124 S1143 S1168 T267 T36 T370 T378 T514 T519 T580 T646 T705 T732 T899 T980 T1098 T1158 Y23 Y29 Y489 Y607	N579	Transmembrane domains: S77-V94, L276-W298, Y330-R350, L947-I971, Q991-I1009	HMMER
					E1-E2 ATPase domains: E381-V403, Q530-A562, Y633-G685, R788-D818	HMMER_PFAM
					E1-E2 ATPases phosphorylation site proteins BL00154: G134-L151, V386-F404, D650-M690, T809-S832	BLIMPS_BLOCKS
					E1-E2 ATPases phosphorylation site: A372-V417	PROFILESCAN
					P-type cation-transporting ATPase superfamily signature PR00119: F390-F404, A666-D676, I812-I831	BLIMPS_PRINTS
					ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATPBINDING PROTEIN PROBABLE CALCIUMTRANSPORTING CALCIUM TRANSPORT PD004657: S846-P1093	BLAST_PRODOM
					FIC1 PROTEIN PD180313: H1039-W1165	BLAST_PRODOM
					do ATPASE; CALCIUM; TRANSPORTING; DM02405|P32660|318-1225: W128-F418, E466-N910	BLAST_DOMO
					ATPase E1-E2 motif: D392-T398	MOTIFS
5	7473347CD1	467	S149 S175 S344 S37 S390 S411 S419 S427 S53 S96 T100 T136 T157 T355 T356 T366 T41	N126 N197 N220	Transmembrane domain: V332-V351	HMMER

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
5					Neurotransmitter-gated ion-channel domain: P58-Q362, H441-W463	HMMER_PFAM
					Neurotransmitter-gated ion channels signature BL00236: V85-P122, I139-H148, D169-Y207, Y254-A295	BLIMPS_BLOCKS
					Neurotransmitter-gated ion-channels signature: L164-H218	PROFILESCAN
					Neurotransmitter-gated ion-channels signature PR00252: T105-F121, K138-S149, C184-C198, S261-P273	BLIMPS_PRINTS
					Gamma-aminobutyric acid A (GABAA) receptor signature PR00253: F270-W290, V296-V317, V330-V351, Y446-Y466	BLIMPS_PRINTS
					CHANNEL IONIC TRANSMEMBRANE GLYCOPROTEIN POSTSYNAPTIC MEMBRANE RECEPTOR PRECURSOR SIGNAL PROTEIN PD000153: E62-S427	BLAST_PRODOM
					NEUROTRANSMITTER-GATED ION-CHANNELS DM00560|P50573|34-464: S37-V467	BLAST_DOMO
					Neurotransmitter-gated ion channels motif: C184-C198	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
6	7474240CD1	1196	S174 S187 S209 S211 S239 S269 S274 S275 S317 S349 S354 S514 S55 S609 S639 S821 S869 S879 S883 S896 S899 S906 S922 S923 S939 S940 S963 S974 S985 S1020 S1091 S1170 S1096 T133 T169 T344 T371 T392 T528 T582 T637 T673 T74 T829 T857 T916 T1022 T1027 T1134 T1099 Y248 Y446 Y98	N102 N230 N338 N369 N600 N661 N736 N881 N905 N1139	Transmembrane domain: V551-Y571	HMMER
					Transmembrane region cyclic nucleotide gated ion channel: Y492-I731	HMMER_PFAM
					Cyclic nucleotide-binding domain: M759-E850	HMMER_PFAM
					POTASSIUM CHANNEL IONIC CHANNEL PD104127: S852-Y1028	BLAST_PRODOM
					POTASSIUM CHANNEL IONIC CHANNEL PD104126: A1076-K1196	BLAST_PRODOM
					CAMP RECEPTOR PROTEIN CYCLIC NUCLEOTIDE-BINDING DOMAIN DM01165|I38465|562-948: H564-A914	BLAST_DOMO
					do POTASSIUM; CHANNEL; KST1; AKT1; DM02383|I38465|353-560: S353-A563	BLAST_DOMO
					do CHANNEL; POTASSIUM; EAG; DM05484|I38465|1-351: M1-P351	BLAST_DOMO
7	7475338CD1	512	S222 S279 S412 S413 S438 T107 T170 T235 T247 T473 T59 T66 Y380	N41 N57	Signal peptide: M1-A35	SPSCAN
					Transmembrane domains: C79-G96, M171-L188, Y322-V342, F448-I466	HMMER
					Sugar (and other) transporter domain: A26-F481	HMMER_PFAM
					Sugar transport proteins signatures: A119-I185, V323-S379	PROFILESCAN
					Sugar transporter signature PR00171: A35-V45, V135-M154, Q294-Y304, I383-V404, T406-F418	BLIMPS_PRINTS
					Glucose transporter signature PR00172: L284-Y305, Q321-V342, L352-Q372, I383-T406, A416-F434, Y446-I466	BLIMPS_PRINTS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
7					SUGAR TRANSPORT PROTEINS DM00135|P22732|132-466: R138-T473	BLAST_DOMO
					Sugar transporter 1 motif: S338-A353	MOTIFS
					Sugar transporter 2 motif: V140-R165	MOTIFS
8	7476747CD1	568	S143 S365 S4 S456 S46 S51 S55 T34 T430 Y45	N141 N205 N214 N256 N562 N62 N76	Transmembrane domains: I242-F269, Y289-P308, I322-Y342	HMMER
					Transmembrane amino acid transporter protein domain: A102-G543	HMMER_PFAM
					ACID AMINO PROTEIN TRANSPORTER PERMEASE TRANSMEMBRANE INTERGENIC REGION PUTATIVE PROLINE PD001875: W80-L380	BLAST_PRODOM
9	7477898CD1	958	S105 S140 S145 S200 S26 S283 S288 S458 S488 S55 S670 S706 S724 S751 S774 S788 S864 S872 S879 S897 S929 T13 T170 T202 T220 T301 T326 T363 T377 T486 T522 T678	N218 N449 N510 N742	Transmembrane domain: L300-N318	HMMER
					Transmembrane region cyclic nucleotide gated ion channel: Y341-I580	HMMER_PFAM
					Cyclic nucleotide-binding domain: V608-A699	HMMER_PFAM
					POTASSIUM CHANNEL IONIC CHANNEL PD118772: E702-S955	BLAST_PRODOM
					CHANNEL PROTEIN IONIC POTASSIUM NONPHOTOTROPIC HYPOCOTYL PUTATIVE SUBUNIT REPEAT EAG PD009483: M1-L86	BLAST_PRODOM
					CAMP RECEPTOR PROTEIN CYCLIC NUCLEOTIDE-BINDING DOMAIN DM01165|I38465|562-948: H413-F738,	BLAST_DOMO
					do POTASSIUM; CHANNEL; KST1; AKT1; DM02383|I38465|353-560: T201-A412	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
10	7472728CD1	724	S229 S283 S303 S333 S512 S545 S597 S666 S718 T104 T19 T223 T444 T515 T540 T557 T591 T636 T640 T650 T661 T676	N327 N330 N331 N532 N664 N684 N716	Transmembrane domains: A370-L388, I419-F437, V486-M503	HMMER
					TASK K+ channel domain: M250-D646	HMMER_PFAM
					TWIK1 RELATED POTASSIUM CHANNEL, SUBFAMILY K, MEMBER 2 TREK1 K+ CHANNEL SUBUNIT IONIC CHANNEL PD085853: P215-G326	BLAST_PRODOM
11	7474322CD1	470	S134 S142 S245 S326 S355 S408 S411 S415 S432 S452 T15 T22 T229 T265 T337 T341 T36	N236 N256 N321 N380	Transmembrane domains: F62-Y87, F139-F163, F212-L230, I293-I312	HMMER
					VANILLOID RECEPTOR SUBTYPE 1 PD137334: C348-K470	BLAST_PRODOM
12	5455621CD1	618	S110 S265 S313 S373 S490 S550 S565 S576 S594 T154 T237 T268 T360 T37 T526 T567 T70	N219 N256 N480 N574	Transmembrane domains: D10-F28, F81-Y104, F278-M297, L439-Y459, I502-R528	HMMER
					Sodium:solute symporter family domain: F41-G445	HMMER_PFAM
					Sodium:solute symporter signature BL00456: T154-G208	BLIMPS_BLOCKS
					Sodium:solute symporter family signature: N151-T198	PROFILESCAN
					TRANSMEMBRANE TRANSPORT PERMEASE PROTEIN SODIUM SYMPORT PROLINE COTRANSPORTER SYMPORTER GLYCOPROTEIN PD000991: F41-C304	BLAST_PRODOM
					SYMPORTER SODIUM IODIDE THYROID SODIUM/IODIDE NIS PD024705: I446-L489, S490-G575	BLAST_PRODOM
					SODIUM:SOLUTE SYMPORTER FAMILY DM00745|P31636|24-561: D10-N219, G220-Y459	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
13	7477248CD1	631	S149 S212 S258 S522 S9 T518 T551 T73 T79 Y14	N352 N516 N96	Transmembrane domains: V22-F41, L159-M181, I391-A407	HMMER
					Sodium/hydrogen exchanger family domain: L25-V491	HMMER_PFAM
					Na+/H+ exchanger isoform 6 signature PR01088: Y14-I38, W39-V57, Y58-V84, Q119-E132, A269-M288, T480-Q506, K515-D533, P539-Q567, P566-E593	BLIMPS_PRINTS
					Na+/H+ exchanger signature PR01084: I133-F144, G147-S161, I162-T170, G208-T218	BLIMPS_PRINTS
					+ TRANSPORT EXCHANGER NA PD01672: I133-M181	BLIMPS_PRODOM
					NA+/H+ PROTEIN TRANSMEMBRANE TRANSPORT ANTIPORTER SYMPORT SODIUM EXCHANGER GLYCOPROTEIN SODIUM/HYDROGEN PD000631: G20-G63, E132-R490	BLAST_PRODOM
					SODIUMHYDROGEN EXCHANGER 6 MYELOBLAST KIAA0267 PD177855: G478-Y591	BLAST_PRODOM
					do BETA; EXCHANGER; NA; DM02572|P48764|10-734: L124-L541	BLAST_DOMO
14	2944004CD1	1256	S103 S130 S144 S170 S227 S252 S523 S802 S817 S899 S901 S98 S1055 T269 T353 T358 T387 T502 T549 T576 T74 T912 T1212 T1061 T1236 Y349 Y407	N150 N23 N300 N312 N318 N704 N1045 N1053 N1059 N1073 N1247	Transmembrane domains: Y231-Y251, L415-L434, I933-I959, F966-L985, I1002-F1020, N1104-M1122	HMMER
					E1-E2 ATPase domains: V274-V365, G490-D506, Q672-A785, L851-S899	HMMER_PFAM
					E1-E2 ATPases phosphorylation site signature BL00154: V454-G490, L492-L510, K652-C662, N724-M764, V878-S901, A905-V938	BLIMPS_BLOCKS
					E1-E2 ATPases phosphorylation site: I478-E526	PROFILESCAN

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
14					P-type cation-transporting ATPase superfamily signature PR00119: N318-T332, C496-L510, A740-D750, C881-L900	BLIMPS_PRINTS
					ATPASE PROBABLE CALCIUMTRANSPORTING PROTEIN HYDROLASE CALCIUM TRANSPORT. TRANSMEMBRANE PHOSPHORYLATION MAGNESIUM PD090368: Q995-Y1094, D1064-L1114	BLAST_PRODOM
					E1-E2 ATPASES PHOSPHORYLATION SITE DM00115|P22189|49-801: S202-K331, P401-E505, S556-A575, V623-P767, H800-S984	BLAST_DOMO
					E1-E2 ATPase motif: D498-T504	MOTIFS
15	3046849CD1	499	S100 S118 S215 S285 T466 T487	N292 N34 N50	Signal peptide: M1-G27	SPSCAN
					Transmembrane domains: M163-L181, T371-G389, M418-L440	HMMER
					Sugar (and other) transporter signature: L18-L474	HMMER_PFAM
					Sugar transport proteins signature: A112-V178	PROFILESCAN
					Sugar transporter signature PR00171: T28-I38, M128-M147, M376-L397, T399-C411	BLIMPS_PRINTS
					Glucose transporter signature PR00172: Q314-I335, M376-T399, A409-L427	BLIMPS_PRINTS
					SUGAR TRANSPORT PROTEINS DM00135|P22732|132-466: R131-T466	BLAST_DOMO
					Sugar transporter 2 motif: L133-R158	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
16	4538363CD1	596	S17 S290 S39 S5 T119 T211	N239 N386 N4 N545 N96	Transmembrane domains: S73-W95, I185-I212, L356-A376, L410-V430, F473-F491, Y513-L533	HMMER
					Sodium:solute *symporter family* domain: Y50-G479	HMMER_PFAM
					Sodium:solute symporter signature BL00456: Y27-G81, A103-R132, L165-G219, P452-G461	BLIMPS_BLOCKS
					Sodium:solute symporter family signatures: H162-I209, V412-D502	PROFILESCAN
					TRANSMEMBRANE TRANSPORT PERMEASE PROTEIN SODIUM SYMPORT PROLINE COTRANSPORTER-SYMPORTER GLYCOPROTEIN PD000991: Y50-G479	BLAST_PRODOM
					NA+/GLUCOSE COTRANSPORTERRELATED PROTEIN PD134393: L551-A596	BLAST_PRODOM
					NA+/GLUCOSE COTRANSPORTERRELATED PROTEIN PD166538: M1-G49	BLAST_PRODOM
					SODIUM:SOLUTE SYMPORTER FAMILY DM00745|P13866|24-561: S17-W548	BLAST_DOMO
					Na solute symporter 2 motif: G461-V481	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
17	6427460CD1	1192	S143 S169 S188 S283 S287 S335 S451 S507 S508 S52 S555 S561 S722 S933 T203 T255 T259 T269 T333 T380 T413 T418 T659 T708 T714 T715 T910 T1103 T1017 T1105 Y885 Y1026	N397 N745 N921 N989 N1001	Transmembrane domains: V299-Y316, F1004-L1022, I1030-W1049, A1075-L1092	HMMER
					E1-E2 ATPase domains: E403-E425 I550-C698	HMMER_PFAM
					E1-E2 ATPases phosphorylation site signature BL00154: G149-F166, V408-F426, D663-L703	BLIMPS_BLOCKS
					E1-E2 ATPases phosphorylation site: L395-C442	PROFILESCAN
					P-type cation-transporting ATPase superfamily signature PR00119: F412-F426, A679-D689	BLIMPS_PRINTS
					ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATPBINDING PROTEIN PROBABLE CALCIUMTRANSPORTING CALCIUM TRANSPORT PD004657: A857-V1108	BLAST_PRODOM
					do ATPASE; CALCIUM; TRANSPORTING; DM02405|Q09891|206-1107: T105-Y436, F471-N921	BLAST_DOMO
					E1-E2 ATPase motif: D414-T420	MOTIFS
18	7474127CD1	638	S205 S224 S336 S378 S414 S541 S553 S564 S86 T120 T146 T155 T17 T21 T25 T283 T374 T49 T520 T546 T579	N259 N266 N518 N536 N84	Transmembrane domains: I231-L248, F382-Y401, M451-V473	HMMER
					Ion transport protein domain: L240-I472	HMMER_PFAM
					Potassium channel signature PR00169: E101-T120, P222-T250, Y284-K307, F310-V330, F352-S378, E381-E404, F421-M443, G450-F476	BLIMPS_PRINTS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
18					VOLTAGEGATED POTASSIUM CHANNEL PROTEIN KV3.2 KSHIIIA IONIC TRANSMEMBRANE ION TRANSPORT GLYCOPROTEIN MULTIGENE FAMILY ALTERNATIVE SPLICING PHOSPHORYLATION PD085814: K495-S538	BLAST_PRODOM
					do CHANNEL; POTASSIUM; CDRK; FORM; DM00436|P22462|189-350: R189-R351	BLAST_DOMO
					do CHANNEL; POTASSIUM; CDRK; SHAW; DM00490|P22462|34-151: L34-C152	BLAST_DOMO
19	7476949CD1	681	S307 S421 S56 S573 S582 S587 S638 S651 T422 T485 T650 Y510	N113 N251 N256 N403 N603	Transmembrane domains: I38-I57, S90-W112, I150-I167, L188-M207, L373-A393, V432-I448, Y530-L550	HMMER
					Sodium:solute symporter family domain: Y67-G496	HMMER_PFAM
					Sodium:solute symporter signature BL00456: Y44-G98, A120-R149, L182-G236, P469-A478	BLIMPS_BLOCKS
					Sodium:solute symporter family signatures: Q179-V226, D458-D519	PROFILESCAN
					TRANSMEMBRANE TRANSPORT PERMEASE PROTEIN SODIUM SYMPORT PROLINE COTRANSPORTER SYMPORTER GLYCOPROTEIN PD000991: Y67-G496	BLAST_PRODOM
					SODIUM:SOLUTE SYMPORTER FAMILY DM00745|P13866|24-561: H34-W565	BLAST_DOMO
					Na solute symporter 1 motif: G183-A208	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
20	7477249CD1	1096	S115 S163 S276 S280 S332 S333 S404 S454 S46 S461 S462 S508 S514 S671 S863 S891 S1084 T262 T340 T345 T347 T407 T570 T612 T687 T840 T948 T1034 T1036 Y322	N331 N383 N395 N411 N720 N932	Transmembrane domains: F289-L307, F935-L953, W967-V996, F1008-D1028	HMMER
					E1-E2 ATPase domains: T340-Q352, H502-V648	HMMER_PFAM
					E1-E2 ATPases phosphorylation site signature BL00154: G143-L160, V335-F353, K529-C539, D616-H656	BLIMPS_BLOCKS
					P-type cation-transporting ATPase superfamily signature PR00119: F339-F353, A632-D642	BLIMPS_PRINTS
					H+-transporting ATPase signatur PR00120: T547-A565	BLIMPS_PRINTS
					ATPASE HYDROLASE TRANSMEMBRANE PHOSPHORYLATION ATPBINDING PROTEIN PROBABLE CALCIUMTRANSPORTING CALCIUM TRANSPORT PD004657: A787-K1038	BLAST_PRODOM
					do ATPASE; CALCIUM; TRANSPORTING; DM02405|P39524|236-1049: T83-I306, F422-N851	BLAST_DOMO
					E1-E2 ATPase motif: D341-T347	MOTIFS
21	7477720CD1	707	S204 S299 S360 S417 S488 S51 S58 S585 S591 S620 S638 S679 T334 T350 T483 T634 Y225 Y528	N297 N31 N342 N35	Signal peptide: M1-A26	SPSCAN
					Transmembrane domains: I155-Y178, I271-T292,	HMMER
					Sodium/hydrogen exchanger family domain: V73-K482	HMMER_PFAM
					Na+/H+ exchanger signature PR01084: I158-A166, G200-A210, I129-L140, G143-S157	BLIMPS_PRINTS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
21					Na+/H+ exchanger isoform 2 (NHE2) signature PR01086: F115-S128, K616-I627	BLIMPS_PRINTS
					+ TRANSPORT EXCHANGER NA PD01672: A83-I113, I129-L177, Y178-L212, A213-F249, D262-I287, S288-Y321, L322-M355, S359-F405, Y406-F452, I489-K531, I532-G562, R593-R640	BLIMPS_PRODOM
					NA+/H+ PROTEIN TRANSMEMBRANE TRANSPORT ANTIPORTER SYMPORT SODIUM EXCHANGER GLYCOPROTEIN SODIUM/HYDROGEN PD000631: I77-A438	BLAST_PRODOM
					do BETA; EXCHANGER; NA; DM02572|P26434|14-716: L15-L687	BLAST_DOMO
22	7477852CD1	729	S142 S144 S155 S285 S291 S299 S318 S654 S664 S669 S697 S719 T110 T138 T281 T379 T447 T532 T539	N208 N358 N717	Transmembrane domains: F493-F512, M554-M570	HMMER
					Ankyrin repeats: L78-E108, A116-T148, F162-S194	HMMER_PFAM
					VANILLOID RECEPTOR SUBTYPE 1 PD101189: F115-L220	BLAST_PRODOM
					ATP/GTP binding site (P-loop): A412-T419	MOTIFS
23	1471717CD1	492	S13 S18 S225 S314 S373 T323 T33 T351 T426	N229 N249	transmembrane domain: I48-V71, V86-F104, Y172-I199, I199-V217, F384-F402, V452-C472	HMMER
					Sugar (and other) transporter: I48-K492	HMMER_PFAM
					SUGAR TRANSPORT PROTEINS DM00032|P30638|80-152:R45-K115	BLAST_DOMO
					VESICLE; SYNAPTIC; SV2; FORM DM08835|S34961|180-344:I119-N249	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
24	3874406CD1	1494	S30 S50 S134 S230 S368 S549 S638 S669 S686 S696 S792 S800 S831 S912 S1004 S1070 S1146 S1172 S1206 S1365 T111 T435 T449 T501 T520 T632 T649 T657 T729 T845 T1049 T1134 T1217 T1247 T1295 T1318 T1339 T1422 T1482 Y824	N109 N130 N313 N421 N453 N71 N788 N817 N84 N867 N91 N1182	transmembrane domain: L204-F221, T272-L290, L735-Y753, F896-S914, V941-I959, L975-R998, F1019-V1039	HMMER
					ABC transporter: G384-G566 G1190-G1366	HMMER_PFAM
					ABC transporters family proteins BL00211: I389-L400, L492-D523	BLIMPS_BLOCKS
					ABC transporters family signature: V472-D523	PROFILESCAN
					ABC TRANSPORTERS FAMILY DM00008|P41233|839-1045:I355-N565, K1177-M1363 DM00008|P34358|611-816:I355-N565, A1179-M1363 DM00008|P41233|1851-2058:K1173-S1365, I355-N565 DM00008|P23703|41-246:E1162-G1366, L377-G566	BLAST_DOMO
					ATP/GTP-binding site motif A (P-loop): G391-S398, G1197-2004	MOTIFS

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
25	4599654CD1	774	S355 S356 S40 S505 S552 S559 S597 S61 S67 S734 S736 T203 T418 T668 T764 Y490	N291 N416	transmembrane domain: Y95-F118, T203-L219, L327-L353	HMMER
					Transmembrane region cyclic Nucleotide G: Y168-I414	HMMER_PFAM
					Cyclic nucleotide-binding domain: K443-M531	HMMER_PFAM
					Cyclic nucleotide-binding domain proteins BL00888: G452-V475, G488-L497	BLIMPS_BLOCKS
					cAMP-dependent protein kinase signature PR00103: F449-R463, S489-T498	BLIMPS_PRINTS
					HYPERPOLARIZATIONACTIVATED CATION CHANNEL, HAC3 PD180735: T538-M774	BLAST_PRODOM
					CHANNEL IONIC POTASSIUM K+ SUBUNIT HYPERPOLARIZATIONACTIVATED PROTEIN PUTATIVE EAG LONG PD001039: E74-R167	BLAST_PRODOM
					CAMP RECEPTOR PROTEIN CYCLIC NUCLEOTIDE-BINDING DOMAIN DM01165|A55251|333-706:H263-P561 DM01165|P29973|311-684:H263-P561 DM01165|Q03041|286-658:H263-G548 DM01165|S52072|262-635:H263-Q595	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
26	5047435CD1	614	S116 S210 S290 S538 S577 S606 T267 T432 T443 T591	N407 N599	transmembrane domain: V124-I142, A168-M190, A371-V390, W483-I511, S526-I543, F552-V570	HMMER
					Sugar (and other) transporter: L83-F585	HMMER_PFAM
					Sugar transport proteins BL00216: L174-S223, G92-S103	BLIMPS_BLOCKS
					Sugar transporter signature PR00171: G92-I102, V175-I194, L486-V507, S509-F521	BLIMPS_PRINTS
					Glucose transporter signature PR00172: V343-V364, L486-S509, R519-L537, W550-V570	BLIMPS_PRINTS
					Sugar_Transport_1: G138-G153 A360-A375	MOTIFS
					Sugar transport proteins signatures sugar_transport_1.prf: L344-S401 sugar_transport_2.prf: A160-A225	PROFILESCAN
					SUGAR TRANSPORT PROTEINS DM00135|S25015|122-478:A160-D417, L480-K574, DM00135|P09830|101-452:G161-V405, L481-K574 DM00135|Q01440|101-433:R178-G388, R178-G388, L486-G575 DM00135|P15729|242-463:A485-S577, R286-L414	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
27	7475603CD1	2180	S181 S216 S233 S260 S409 S419 S842 S983 S1008 S1172 S1229 S1237 S1269 S1349 S1353 S1462 S1469 S1504 S1566 S1881 S1993 S2018 S2174 S2167 T120 T165 T338 T348 T510 T599 T614 T822 T931 T1079 T1086 T1094 T1171 T1181 T1209 T1219 T1417 T1439 T1822 T1870 T1917 T1988 T2057 T2125 Y656 Y1448	N112 N132 N346 N374 N1100 N1415 N1420 N1491 N1552 N1695 N1831	transmembrane domain: F630-L648, L664-L680, V1570-V1590, M1622-Q1641	HMMER
					ABC transporter: G1854-G2035 G868-G1048	HMMER_PFAM
					ABC transporters family BL00211: F873-T884, L974-D1005	BLIMPS_BLOCKS
					ABC transporters family signature: A1940-D1991, D955-D1005	PROFILESCAN
					Abc_Transporter: L974-F988	MOTIFS
					ATP/GTP-binding site motif A (P-loop): G875-T882, G1861-T1868	MOTIFS
					ATPBINDING TRANSPORTER CASSETTE ABC TRANSPORT PROTEIN GLYCOPROTEIN TRANSMEMBRANE RIM ABCR PD005939: L1563-N1740	BLAST_PRODOM
					ATPBINDING TRANSPORTER CASSETTE ABC GLYCOPROTEIN TRANSMEMBRANE TRANSPORT ABCR RIM PD010118: R238-R514, L95-R243	BLAST_PRODOM
					ATPBINDING TRANSPORTER CASSETTE ABC GLYCOPROTEIN TRANSMEMBRANE TRANSPORT ABCR RIM SIMILARITY PD008845: P1307-E1560	BLAST_PRODOM
					ATPBINDING TRANSPORTER CASSETTE ABC GLYCOPROTEIN TRANSMEMBRANE TRANSPORT RIM ABCR SIMILARITY PD006867: L540-S685, D515-Q541	BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
27					ABC TRANSPORTERS FAMILY DM00008|P41233|839-1045:V841-A1046, L1829-M2032 DM00008|P41233|1851-2058:V1826-N2034, V841-V1045 DM00008|P34358|1441-1640:L1827-M2032, V843-V1045	BLAST_DOMO
28	7477845CD1	1737	S23 S254 S687 S692 S695 S7 S713 S766 S773 S8 S861 S1113 S1228 S1271 S1455 S1463 S1537 S1595 S1647 S1652 S1730 T272 T324 T886 T1257 T1320 T1359 T1387 T1406 T1456 T1486 T1528 T1561 T1570 T1645 T1694 Y419 Y702 Y832	N210 N216 N859 N1064 N1371 N1449	transmembrane domain: M1244-A1262, V1319-F1336, I1338-F1357, A1423-I1446, W107-V126, V181-M199, S298-I321, L509-V531, V575-I598, Y879-M904, I1017-F1034, I1134-V1152	HMMER
					Ion transport protein ion_trans: W32-I321 M380-I598 L884-V1155 I1206-I1446	HMMER_PFAM
					Calcium channel signature PR00167: D535-D561	BLIMPS_PRINTS
					PROTEIN F17C8.6 C11D2.5 NEARLY IDENTICAL C ELEGANS PREDICTED PD023984: V1447-S1637, E1714-T1720	BLAST_PRODOM
					C11D2.6 PROTEIN PD178227: L1241-R1368, I1206-F1292 F585-E606	BLAST_PRODOM
					C11D2.6 PROTEIN SIMILARITY ALONG ENTIRE GENE CALCIUM CHANNEL ALPHA PROTEINS PD041964: L599-V885,	BLAST_PRODOM
					CHANNEL CALCIUM IONIC SUBUNIT VOLTAGE GATED SODIUM ALPHA TRANSMEMBRANE L TYPE PD000032: Y887-V1120, I33-V330, K1361-F1450, I1206-F1357, I577-I598, F1337-L1356, I1134-F1159, D1416-V1443	BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
28					III REPEAT DM00079|A55138|1052-1268:V1020-L1227 DM00079|P35500|1424-1636:W1090-P1194, I1017-N1050	BLAST_DOMO
					IV REPEAT DM00277|P27732|1363-1572:F1337-LI536 DM00277|P15381|1384-1595:F1337-L1536	BLAST_DOMO
29	168827CD1	547	S109 S167 S201 S282 S336 S404 S408 S526 T133 T323 T35 T432 T453 T58	N102 N107 N56	transmembrane domain: F16-T35, Y180-C200, S201-V222, M410-E429, T469-Y492, L496-L514	HMMER
					Sugar (and other) transporter: L13-Q528	HMMER_PFAM
					ORGANIC TRANSPORTERLIKE TRANSPORT PROTEIN RENAL ANION TRANSPORTER CATIONIC KIDNEYSPECIFIC SOLUTE PD151320: N102-L144	BLAST_PRODOM
30	7472734CD1	547	S143 S167 S201 S282 S336 S404 S408 S46 S526 S60 S68 T133 T323 T432 T453 T58	N102 N39 N56 N62	transmembrane domain: I18-F32, M147-Y163, Y180-C200, S201-V222, M410-E429, T469-Y492, L496-L514	HMMER
					Sugar (and other) transporter: L18-Q528	HMMER_PFAM
					SUGAR TRANSPORT PROTEINS DM00032|P46501|280-351:V121-K173	BLAST_DOMO
					ORGANIC TRANSPORTERLIKE TRANSPORT PROTEIN RENAL ANION TRANSPORTER CATIONIC KIDNEYSPECIFIC SOLUTE PD151320: N102-K145	BLAST_PRODOM

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
31	7473473CD1	988	S142 S237 S24 S252 S322 S369 S502 S680 S773 S847 S883 S925 S943 S952 S974 S981 T127 T14 T215 T442 T478 T521 T634 T725 T73 T832 T869 T909 T929	N170 N235 N403 N466 N663 N830	transmembrane domain: L342-A360	HMMER
					Transmembrane cyclic Nucleotide G: Y288-I536	HMMER_PFAM
					Cyclic nucleotide-binding domain: V564-A655	HMMER_PFAM
					PAC motif PA: C92-T132	HMMER_PFAM
					CHANNEL POTASSIUM IONIC EAG SUBUNIT HEAG LONG ELECTOCARDIOGRAPHIC QT SYNDROME PD017645: K809-D984	BLAST_PRODOM
					CHANNEL IONIC K+ SUBUNIT HYPERPOLARIZATION ACTIVATED PUTATIVE EAG LONG PD001039: S179-I284	BLAST_PRODOM
					CHANNEL K+ IONIC EAG SUBUNIT TRANSMEMBRANE ION TRANSPORT VOLTAGEGATED PD011550: N658-E737	BLAST_PRODOM
					CHANNEL PROTEIN IONIC POTASSIUM NON PHOTOTROPIC HYPOCOTYL PUTATIVE SUBUNIT REPEAT EAG PD009483: M1-E89	BLAST_PRODOM
					CAMP RECEPTOR PROTEIN CYCLIC NUCLEOTIDE-BINDING DOMAIN DM01165|I48912|391-786:H361-S756 DM01165|Q02280|384-776:H361-E737 DM01165|I38465|562-948:H361-R671, S974-E985	BLAST_DOMO
					POTASSIUM; CHANNEL; KST1; AKT1; DM02383|I48912|164-389:V162-E314, E314-A360, W362-V455	BLAST_DOMO

Table 3 (cont.)

SEQ ID NO:	Incyte Polypeptide ID	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Domains and Motifs	Analytical Methods and Databases
32	7477725CD1	533	S107 S109 S143 S167 S282 S345 S408 S469 S60 T133 T289 T323 T336 T432 T526	N102 N216 N56 N62	transmembrane domain: F150-D168, L380-N401, I407-V426, L486-F504	HMMER
					Sugar (and other) transporter: A111-K528	HMMER_PFAM
					ORGANIC TRANSPORTER LIKE TRANSPORT PROTEIN RENAL ANION TRANSPORTER CATIONIC KIDNEY SPECIFIC SOLUTE PD151320: N102-K145	BLAST_PRODOM

Table 4

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
33	3474673CB1	1775	1-391, 578-786, 1024-1301	GNFL.g7798848_000003_004.edit	1	1156
				6724643H1 (LUNLTMT01)	861	1347
				3474673H1 (LUNGNOT27)	249	568
				71495515V1	1205	1775
34	4588877CB1	1545	261-619, 1-193, 794-1071	71495515V1)	975	1545
				FL135171_00001	539	1534
				71497982V1	1	662
35	7472214CB1	1941	1483-1558, 1-413, 495-616, 732-1149	GBI:g8117242_000054_edit.8639-8803	1171	1335
				GBI:g8117242_000054_edit.4857-4997	544-	684
				GBI:g8117242_000054.edit.10305-10463	1441	1599
				6891360H1 (BRAITDR03)	1433	1905
				GBI:g8117242_000054_edit.50-89	1	240-
				GBI:g8117242_000054_edit.6950-7093	925	1068
				GBI:g8117242_000054_edit.4345-4478	358	492
				60124962D2	1735	1941
				GBI:g8117242_000054_edit.8313-8414	1069	1170
				GBI:g8118985_000043_edit.12301-12444.comp	685	810
				GBI:g8117242_000054_edit.4112-4228	241	357
				GBI:g8117242_000054_edit.10957-11181	1717	1941
				5500380H1 (BRABDIR01)	907	1119
				GBI:g8117242_000054_edit.10616-10732	1600	1716

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
35				GBI:g8117242_000054_edit.8907-9011	1336	1440
				GBI:g8117242_000054_edit.6643-6756	811	924
36	7473053CB1	4971	3312-3482, 1-1466, 4307-4971, 2184-2221	8035016H1 (SMCRUNE01)	2315	2975
				6822202J1 (SINTNOR01)	2145	2877
				6781747H1 (OVARDIR01)	968	1449
				8035016J1 (SMCRUNE01)	2979	3643
				6824230H1 (SINTNOR01)	2867	3483
				6894266H1 (BRAITDR03)	548	1157
				6777836H1 (OVARDIR01)	1601	2238
				6908503H1 (PITUDIR01)	1	667
				6908503J1 (PITUDIR01)	1270	1830
				6823447H1 (SINTNOR01)	3525	4260
				6823447J1 (SINTNOR01)	4226	4829
				6006310F8 (FIBRUNT02)	4501	4969
				4171959T6 (SINTNOT21)	3637	4287
				5088860F6 (UTRSTMR01)	4461	4853
37	7473347CB1	1404	126-633, 1013-1404, 768-838	GBI.lee4.edit	1	1404

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
38	7474240CB1	4048	3023-4048, 1753-2469, 1-920, 1593-1658, 2614-2908, 1138-1367	71984804V1	964	1311
				GBI:g7656646_edit	929	3418
				71986624V1	1369	1976
				55055014H1	1	130
				55037111J2	95	871
				71983668V1	1371	2043
				GBI:g5923734_edit	2612	4048
				55037119J2	224	875
				2502027F6 (ADRETUT05)	696	1235
39	7475338CB1	1539	1412-1539, 1-328, 495-837, 922-1218	GBI:g7960701_000004_edit.549-713	154	312
				GBI:g7960701_000004_edit.13381-13480	1015	1113
				GBI:g7960701_000004_edit.8755-8943	715	903
				GBI:g7960701_000004_edit.4292-4417	313	438
				GBI:g7960701_000004_edit.16237-16317	1114	1194
				GBI:g7960701_000004_edit.20107-20325	1321	1539
				GBI:g7960701_000004_edit.9989-10099	904	1014
				GBI:g7960701_000004_edit.18748-18873	1195	1320
				GBI:g7960701_000003_edit.9783-9884	52	153
				GBI:g7960701_000004_edit.5251-5403	439	591
				GBI:g7960701_000004_edit.8384-8506	592	714
				71906448V1	627	1082
				71753467V1	912	1539
40	7476747CB1	3114	1717-1870, 1-503, 1468-1650	3351512F6 (PROSNOT28)	2185	2724
				7761783J1 (THYMNOE02)	1943	2570
				6934981R8 (SINTTMR02)	78	860

Table 4 (cont.)

Polynucleotide *SEQ ID NO:*	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5′ Position	3′ Position
40				6389368H1 (PROSTMC01)	1782	2075
				70536163V1	2575	3114
				6934981F8 (SINTTMR02)	1	643
				GNN.g7712065_000012_002	452	1922
				7080657H1 (STOMTMR02)	838	1403
				5633289H1 (PLACFER01)	639	890
				g5746200	1215	1473
41	7477898CB1	2877	846-901, 1272-1378, 2319-2877	GBI.g2262095	1	2877
42	7472728CB1	2820	1-1399, 2207-2229	55022826J1	1138	1834
				55030210H1	403	986
				4399366T6 (TESTTUT03)	2231	2777
				55030274H1	1482	2153
				g565876	2597	2820
				55018149J1	1907	2585
				FL203597_00001	712	1807
				GNN.g7263861_026.edit	1	1052
43	7474322CB1	1440	1-604, 714-768	GBI.g8081632_edit	1	1440
				71228887V1	1090	1440
				70868623V1	988	1385
44	5455621CB1	2394	1483-1686, 1-329, 838-1155, 2201-2235	3696546T6 (SININOT05)	1833	2394
				70674954V1	1520	2091
				1426382H1 (SINTBST01)	1224	1492
				3696546F6 (SININOT05)	799	1381
				6828352H1 (SINTNOR01)	530	1149
				3699565H1 (SININOT05)	1	281
				7700096H1 (KIDPTDE01)	250	990
				70678552V1	1419	2055

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
45	7477248CB1	2890	1-58, 2739-2890, 2310-2349, 329-1167	2777287H1 (OVARTUT03)	2250	2498
				7977733H1 (LSUBDMC01)	841	1427
				7678168J1 (NOSETUE01)	1271	1827
				7611941J1 (KIDCTME01)	2273	2890
				6590507H1 (TLYMUNT03)	179	672
				2701794F6 (OVARTUT10)	1208	1741
				2544096F6 (UTRSNOT11)	1732	2252
				60117044D2	1	431
				5020832H1 (OVARNON03)	2195	2471
				7662529H1 (UTRSTME01)	526	926
46	2944004CB1	3926	3338-3365, 1-687, 1222-2267	4762728F6 (PLACNOT05)	872	1387
				g2264624	2268	2446
				6264977H1 (MCLDTXN03)	1210	1797
				2944004F6 (BRAITUT23)	2790	3531
				6610392H2 (MUSTTMC01)	3306	3926
				GNN.g7328818_000024_002.edit	2145	2648
				7035078H1 (SINTFER03)	1	440
				7620248J1 (HEARFEE03)	2431	3039
				496537H1 (HNT2NOT01)	2329	2487
				6264427T8 (MCLDTXN03)	453	1174
				6264427F8 (MCLDTXN03)	170	842

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
46				7673654H1 (FIBPFEC01)	1733	2239
47	3046849CB1	2135	2072-2135, 596-711, 1014-1263	8262790U1	1383	2135
				71896642V1	1	592
				71247870V1	1050	1736
				FL3046849_g6815043_000004_g183298	51	1520
48	4538363CB1	2637	1-183, 1575-1680, 2094-2637	FL4538363_g3126781_g520469	1	1917
				71401405V1	1766	2637
49	6427460CB1	3783	985-1833, 2687-3204	70857895V1	416	1035
				7727961J1 (UTRCDIE01)	3284	3783
				70857789V1	566	1109
				g5689372_edit	1092	3361
				g3801917	1	452
50	7474127CB1	2105	1078-2105	GBI.g8568959_edit_3	1119	2105
				g6140313	482	951
				5819744F7 (PROSTUS23)	168	479
				g5920552	1	488
				55049678J1	862	1359
51	7476949CB1	2069	1233-1356, 1-117, 2047-2069, 347-503, 1536-1844	FL7476949_g6714723_g338053	1	2046
				4669722H1 (SINTNOT24)	1801	2069
52	7477249CB1	4245	2833-3018, 1869-2121, 3707-4245, 1-252, 982-1239, 289-357	71660072V1	2404	3156
				71657569V1	3106	3854
				7633968J1 (SINTDIE01)	2579	3175
				6440145F8 (BRAENOT02)	938	1087
				71664080V1	3228	3891
				GBI.g8567478.edit	1	2547
				71660176V1	3773	4245
				71662066V1	1802	2475
				2605539F6 (LUNGTUT07)	433	939
				71659261V1	1690	2437
				3825558H1 (BRAIHCT02)	1179	1270

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
52				7765571H1 (URETTUE01)	1	693
				5675861H1	1427	1716
53	7477720CB1	2124	1-936, 1200-1488, 1982-2124, 1562-1745	FL7477720_g5836195_g205709	1	2124
54	7477852CB1	2195	1-418, 1899-2195	GBI.g8748866.edit	1	2195
55	1471717CB1	2055	206-768, 881-931, 1155-1323	70464956V1	492	994
				72277206V1	1	297
				70469664V1	939	1582
				GNN.g7109510_000068_002.edit	772	1500
				GBI.g8039708_50_63_62_56.edit	238	897
				6540941H1 (LNODNON02)	1571	2055
				70466394V1	1035	1616
56	3874406CB1	4727	1-1299, 1576-1632, 2550-3619, 2014-2192	71793833V1	4117	4727
				55052105J1	1673	2128
				71798347V1	3620	4358
				71798870V1	3575	4244
				55058313J1	1380	2125
				55051482J1	2475	3134
				FL3874406_g3810670_g4240130_3_3-4	482	744
				55068154H1	2223	2741
				3133035F6 (SMCCNOT01)	1	605
				55058329H1	723	1528
				55068182J1	2048	2685
				71795307V1	2902	3593
57	4599654CB1	3852	1-335, 2014-3231	8016331J1 (BMARTXE01)	1778	2424
				71040001V1	3348	3852
				8041905H1 (OVARTUE01)	1666	2352
				55062505H1	660	1233
				g7959336_CD	349	2540
				6772024J1 (BRAUNOR01)	1	623
				55064208J1	1118	1718

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
57				6617183H2 (BRAXTDR14)	2981	3530
				6195941H1 (PITUNON01)	2823	3458
				71909238V1	1225	1747
				2216896F6 (SINTFET03)	2474	2923
				71042073V1	2276	2745
58	5047435CB1	1917	1-238, 1162-1474	7431853H1 (UTRMTMR02)	1211	1917
				GNN:g4375937_004_edit	1	1845
				6426880H1 (LUNGNON07)	814	1336
				6781142H1 (OVARDIR01)	224	941
				2645767H1 (OVARNOT09)	128	394
59	7475603CB1	6791	1-3283, 5952-6101, 3793-4761	71704421V1	6240	6791
				7726210H1 (THYRDIE01)	1885	2602
				7721710J2 (THYRDIE01)	2696	3232
				6340173F8 (BRANDIN01)	5516	6222
				71704256V1	3025	3734
				7757131H1 (SPLNTUE01)	2408	3093
				GNN.g7711543_000002_002.edit	198	2751
				7464813H1 (LIVRFEE04)	544	696
				71703676V1	3250	3947
				7760618H1 (THYMNOE02)	2183	2676
				71970086V1	5817	6525
				7462584H1 (LIVRFEE04)	1	578
				7760618J1 (THYMNOE02)	1251	1983
				71762287V1	4313	4879

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
59				7724639H1 (THYRDIE01)	951	1545
				55052451J1	4792	5698
				7739867H1 (THYMNOE01)	5131	5794
				6879936H1 (UTRSTMR02)	697	1054
				55058371H1	3850	4747
60	7477845CB1	5214	2390-4599, 645-1796	GBI.g8346195_edit	1765	5214
				GBI.g8052096_edit	1132	1839
				8104845H1 (MIXDDIE02)	2822	3367
				GBI.g8518014_edit	1	1266
61	168827CB1	1818	1-281, 796-912	g1081430	1036	1525
				168827H1 (LIVRNOT01)	65	406
				55064792J1	1	209
				55072770H1	495	1110
				GNN.g6498074_012.edit	1321	1818
				087510H1 (LIVRNOT01)	314	574
				g751568	1336	1773
62	7472734CB1	2245	1223-1339, 1-710	55055559H1	16	699
				55045003H2	1	697
				g5361744	908	1109
				GBI.g8118965_000015_000006_000001_000010_000003.edit	602	2245
				g751568	1763	2200
63	7473473CB1	3196	1-376, 460-1796	55049235H1	556	1287
				GBI.g8018151_000001.edit	1799	3196
				GBI.g6433826_000001.edit	1172	2052
				55063069J1	1	850
				g669271	1799	2106

120

Table 4 (cont.)

Polynucleotide SEQ ID NO:	Incyte Polynucleotide ID	Sequence Length	Selected Fragment(s)	Sequence Fragments	5' Position	3' Position
64	7477725CB1	1602	1072-1602	7455614H1 (LIVRTUE01)	416	835
				4288148H1 (LIVRDIR01)	112	257
				GBI.g8131631_000007_000005.edit	1	1602
				g2656651	829	1084

Table 5

Polynucleotide SEQ ID NO:	Incyte Project ID	Representative Library
33	3474673CB1	LUNLTMT01
34	4588877CB1	LUNLTMT01
35	7472214CB1	BRAENOT04
36	7473053CB1	SINTNOR01
38	7474240CB1	ADRETUT05
39	7475338CB1	SINTNOT18
40	7476747CB1	SINTTMR02
42	7472728CB1	TESTTUT03
43	7474322CB1	SINTBST01
44	5455621CB1	SININOT05
45	7477248CB1	UTRSNOT11
46	2944004CB1	MCLDTXN03
47	3046849CB1	HNT2AGT01
48	4538363CB1	PANCNOT07
49	6427460CB1	BRAUNOR01
50	7474127CB1	PROSTUS23
51	7476949CB1	COLNTMC01
52	7477249CB1	COLNPOT01
55	1471717CB1	OVARDIT01
56	3874406CB1	LIVRDIR01
57	4599654CB1	LUNGNOT23
58	5047435CB1	OVARDIR01
59	7475603CB1	THYRDIE01
60	7477845CB1	MIXDDIE02
61	168827CB1	LIVRNOT01
64	7477725CB1	LIVRTUE01

Table 6

Library	Vector	Library Description
ADRETUT05	pINCY	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 52-year-old Caucasian female during a unilateral adrenalectomy. Pathology indicated a pheochromocytoma.
BRAENOT04	pINCY	Library was constructed using RNA isolated from inferior parietal cortex tissue removed from the brain of a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver .
BRAUNOR01	pINCY	This random primed library was constructed using RNA isolated from striatum, globus pallidus and posterior putamen tissue removed from an 81-year-old Caucasian female who died from a hemorrhage and ruptured thoracic aorta due to atherosclerosis. Pathology indicated moderate atherosclerosis involving the internal carotids, bilaterally; microscopic infarcts of the frontal cortex and hippocampus; and scattered diffuse amyloid plaques and neurofibrillary tangles, consistent with age. Grossly, the leptomeninges showed only mild thickening and hyalinization along the superior sagittal sinus. The remainder of the leptomeninges was thin and contained some congested blood vessels. Mild atrophy was found mostly in the frontal poles and lobes, and temporal lobes, bilaterally. Microscopically, there were pairs of Alzheimer type II astrocytes within the deep layers of the neocortex. There was increased satellitosis around neurons in the deep gray matter in the middle frontal cortex. The amygdala contained rare diffuse plaques and neurofibrillary tangles. The posterior hippocampus contained a microscopic area of cystic cavitation with hemosiderin-laden macrophages surrounded by reactive gliosis. Patient history included sepsis, cholangitis, post-operative atelectasis, pneumonia CAD, cardiomegaly due to left ventricular hypertrophy, splenomegaly, arteriolonephrosclerosis, nodular colloidal goiter, emphysema, CHF, hypothyroidism, and peripheral vascular disease.
COLNPOT01	pINCY	Library was constructed using RNA isolated from colon polyp tissue removed from a 40-year-old Caucasian female during a total colectomy. Pathology indicated an inflammatory pseudopolyp; this tissue was associated with a focally invasive grade 2 adenocarcinoma and multiple tubuvillous adenomas. Patient history included a benign neoplasm of the bowel.
COLNTMC01	pINCY	This large size-fractionated library was constructed using pooled cDNA from three different donors. cDNA was generated using mRNA isolated from colon epithelium tissue removed from a 13-year-old Caucasian female (donor A) who died from a motor vehicle accident; from ascending colon removed from a 29-year-old female (donor

Table 6 (cont.)

Library	Vector	Library Description
		B); and from colon tissue removed from the appendix of a 37-year-old Black female (donor C) during myomectomy, dilation and curettage, right fimbrial region biopsy, and incidental appendectomy. Pathology for donor B indicated the proximal and distal resection margins of small bowel and colon away from the mass lesion were uninvolved by lymphoma. Pathology for donor C indicated an unremarkable appendix. Pathology for the matched tumor tissue (donor B) indicated malignant lymphoma, small cell, non-cleaved (Burkitt's lymphoma, B-cell phenotype), forming a polypoid mass in the region of the ileocecal valve, associated with intussusception and obstruction clinically. The liver and multiple (3 of 12) ileocecal region lymph nodes were also involved by lymphoma. Pathology for the associated tumor tissue (donor C) indicated multiple uterine leiomyomata. Donor C presented with deficiency anemia, an umbilical hernia, and premenopausal menorrhagia. Patient history included sarcoidosis of the lung.
HNT2AGT01	PBLUESCRIPT	Library was constructed at Stratagene (STR937233), using RNA isolated from the hNT2 cell line derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor. Cells were treated with retinoic acid for 5 weeks and with mitotic inhibitors for two weeks and allowed to mature for an additional 4 weeks in conditioned medium.
LIVRDIR01	pINCY	The library was constructed using RNA isolated from diseased liver tissue removed from a 63-year-old Caucasian female during a liver transplant. Patient history included primary biliary cirrhosis diagnosed in 1989. Serology was positive for anti-mitochondrial antibody.
LIVRNOT01	PBLUESCRIPT	Library was constructed at Stratagene, using RNA isolated from the liver tissue of a 49-year-old male.
LIVRTUE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from liver tumor tissue removed from a 72-year-old Caucasian male during partial hepatectomy. Pathology indicated metastatic grade 2 (of 4) neuroendocrine carcinoma forming a mass. The patient presented with metastatic liver cancer. Patient history included benign hypertension, type I diabetes, prostatic hyperplasia, prostate cancer, alcohol abuse in remission, and tobacco abuse in remission. Previous surgeries included destruction of a pancreatic lesion, closed prostatic biopsy, transurethral prostatectomy, removal of bilateral testes and total splenectomy. Patient medications included Eulexin, Hytrin, Proscar, Ecotrin, and insulin. Family history included atherosclerotic coronary artery disease and acute myocardial infarction in the mother; atherosclerotic coronary artery disease and type II diabetes in the father.
LUNGNOT23	pINCY	Library was constructed using RNA isolated from left lobe lung tissue removed from

Table 6 (cont.)

Library	Vector	Library Description
		a 58-year-old Caucasian male. Pathology for the associated tumor tissue indicated metastatic grade 3 (of 4) osteosarcoma. Patient history included soft tissue cancer, secondary cancer of the lung, prostate cancer, and an acute duodenal ulcer with hemorrhage. Family history included prostate cancer, breast cancer, and acute leukemia.
LUNLTMT01	pINCY	The library was constructed using RNA isolated from right middle lobe lung tissue removed from a 63-year-old Caucasian female during a segmental lung resection. Pathology for the associated tumor tissue indicated grade3 adenocarcinoma in the right lower lobe and right middle lobe that infiltrated the parietal pleural surface. Metastatic grade 3 adenocarcinoma was found in the diaphragm. The lymph nodes contained metastatic grade 3 adenocarcinoma and involved the superior mediastinal and inferior mediastinal lymph nodes. Patient history included hyperlipidemia. Family history included benign hypertension,cerebrovascular disease, breast cancer, and hyperlipidemia.
MCLDTXN03	pINCY	This normalized dendritic cell library was constructed from one million independent clones from a pool of two derived dendritic cell libraries. Starting libraries were constructed using RNA isolated from untreated and treated derived dendritic cells from umbilical cord blood CD34+ precursor cells removed from a male. The cells were derived with granulocyte/macrophage colony stimulating factor (GM-CSF), tumor necrosis factor alpha (TNF alpha), and stem cell factor (SCF). The GM-CSF was added at time 0 at 100 ng/ml, the TNF alpha was added at time 0 at 2.5 ng/ml, and the SCF was added at time 0 at 25 ng/ml. Incubation time was 13 days. The treated cells were then exposed to phorbol myristate acetate (PMA), and Ionomycin. The PMA and Ionomycin were added at 13 days for five hours. The library was normalized in two rounds using conditions adapted from Soares et al., PNAS (1994) 91:9228-9232 and Bonaldo et al., Genome Research (1996) 6:791, except that a significantly longer (48 hours/round) reannealing hybridization was used.
MIXDDIE02	PBK-CMV	This 5' biased random primed library was constructed using pooled cDNA from seven donors. cDNA was generated using mRNA isolated from brain tissue removed from two Caucasian male fetuses who died after 23 weeks gestation from hypoplastic left heart (A) and prematurity (B); from posterior hippocampus from a 55-year-old male who died from COPD (C); from cerebellum, corpus callosum, thalmus and temporal lobe tissue from a 57-year-old Caucasian male who died from a CVA (D); from dentate nucleus and vermis from an 82-year-old Caucasian male who died from a myocardial infarction (E); from pituitary gland from a 74-year-old Caucasian female who died from a myocardial infarction (F) and vermis tissue from a 77-year-old Caucasian female who died from pneumonia (G). For donor C, pathology indicated

Table 6 (cont.)

Library	Vector	Library Description
		mild lateral ventricular enlargement. For donor F, pathology indicated moderate Alzheimer's disease, recent multiple infarctions involving left thalamus, left parietal and occipital lobes (microscopic) and right cerebellum (gross), mild atherosclerosis involving middle cerebral arteries bilaterally and mild cerebral amyloid angiopathy. For donor G, pathology indicated severe Alzheimer's disease, mild atherosclerosis involving the middle cerebral and basilar arteries, and cerebral atrophy consistent with Alzheimer's disease, For donor D, patient history included Huntington's chorea. Donor E was taking nitroglycerin and dopamine; donor F was taking Lopressor, heparin, ceftriaxone, captopril, Isordil, nitroglycerin, Clinoril, Ecotrin and tacrine; and donor G was taking insulin.
OVARDIR01	PCDNA2.1	This random primed library was constructed using RNA isolated from right ovary tissue removed from a 45-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, vaginal suspension and fixation, and incidental appendectomy. Pathology indicated stromal hyperthecosis of the right and left ovaries. Pathology for the matched tumor tissue indicated a dermoid cyst (benign cystic teratoma) in the left ovary. Multiple (3) intramural leiomyomata were identified. The cervix showed squamous metaplasia. Patient history included metrorrhagia, female stress incontinence, alopecia, depressive disorder, pneumonia, normal delivery, and deficiency anemia. Family history included benign hypertension, atherosclerotic coronary artery disease, hyperlipidemia, and primary tuberculous complex.
OVARDIT01	pINCY	Library was constructed using RNA isolated from diseased ovary tissue removed from a 39-year-old Caucasian female during total abdominal hysterectomy, bilateral salpingo-oophorectomy, dilation and curettage, partial colectomy, incidental appendectomy, and temporary colostomy. Pathology indicated the right and left adnexa were extensively involved by endometriosis. Endometriosis also involved the anterior and posterior serosal surfaces of the uterus and the cul-de-sac and the mesentery and muscularis propria of the sigmoid colon. Pathology for the associated tumor tissue indicated multiple (3 intramural, 1 subserosal) leiomyomata. Family history included hyperlipidemia, benign hypertension, atherosclerotic coronary artery disease, depressive disorder, brain cancer, and type II diabetes.
PANCNOT07	pINCY	Library was constructed using RNA isolated from the pancreatic tissue of a Caucasian male fetus, who died at 23 weeks' gestation.
PROSTUS23	pINCY	This subtracted prostate tumor library was constructed using 10 million clones from a pooled prostate tumor library that was subjected to 2 rounds of subtractive hybridization with 10 million clones from a pooled prostate tissue library. The

Table 6 (cont.)

Library	Vector	Library Description
		starting library for subtraction was constructed by pooling equal numbers of clones from 4 prostate tumor libraries using mRNA isolated from prostate tumor removed from Caucasian males at ages 58 (A), 61 (B), 66 (C), and 68 (D) during prostatectomy with lymph node excision. Pathology indicated adenocarcinoma in all donors. History included elevated PSA, induration and tobacco abuse in donor A; elevated PSA, induration, prostate hyperplasia, renal failure, osteoarthritis, renal artery stenosis, benign HTN, thrombocytopenia, hyperlipidemia, tobacco/alcohol abuse and hepatitis C (carrier) in donor B; elevated PSA, induration, and tobacco abuse in donor C; and elevated PSA, induration, hypercholesterolemia, and kidney calculus in donor D. The hybridization probe for subtraction was constructed by pooling equal numbers of cDNA clones from 3 prostate tissue libraries derived from prostate tissue, prostate epithelial cells, and fibroblasts from prostate stroma from 3 different donors. Subtractive hybridization conditions were based on the methodologies of Swaroop et al., NAR 19 (1991):1954 and Bonaldo, et al. Genome Research 6 (1996):791.
SININOT05	pINCY	Library was constructed using RNA isolated from ileum tissue obtained from a 30-year-old Caucasian female during partial colectomy, open liver biopsy, incidental appendectomy, and permanent colostomy. Patient history included endometriosis. Family history included hyperlipidemia , anxiety, and upper lobe lung cancer, stomach cancer, liver cancer, and cirrhosis.
SINTBST01	pINCY	Library was constructed using RNA isolated from the ileum tissue of an 18-year-old Caucasian female. The ileum tissue, along with the cecum and appendix, were removed during bowel anastomosis. Pathology indicated Crohn's disease of the ileum, involving 15 cm of the small bowel. The cecum and appendix were unremarkable, and the margins were uninvolved. The patient presented with abdominal pain and regional enteritis. Patient history included osteoporosis of the vertebra and abnormal blood chemistry. Patient medications included Prilosec (omeprazole), Pentasa (mesalamine), amoxicillin, and multivitamins. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
SINTNOR01	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 31-year-old Caucasian female during Roux-en-Y gastric bypass. Patient history included clinical obesity.
SINTNOT18	pINCY	Library was constructed using RNA isolated from small intestine tissue obtained from a 59-year-old male.
SINTTMR02	PCDNA2.1	This random primed library was constructed using RNA isolated from small intestine tissue removed from a 59-year-old male. Pathology for the matched tumor tissue

Table 6 (cont.)

Library	Vector	Library Description
		indicated multiple (9) carcinoid tumors, grade 1, in the small bowel. The largest tumor was associated with a large mesenteric mass. Multiple convoluted segments of bowel were adhered to the tumor. A single (1 of 13) regional lymph node was positive for malignancy. The peritoneal biopsy indicated focal fat necrosis.
TESTTUT03	pINCY	Library was constructed using RNA isolated from right testicular tumor tissue removed from a 45-year-old Caucasian male during a unilateral orchiectomy. Pathology indicated seminoma. Patient history included hyperlipidemia and stomach ulcer. Family history included cerebrovascular disease, skin cancer, hyperlipidemia, acute myocardial infarction, and atherosclerotic coronary artery disease.
THYRDIE01	PCDNA2.1	This 5' biased random primed library was constructed using RNA isolated from diseased thyroid tissue removed from a 22-year-old Caucasian female during closed thyroid biopsy, partial thyroidectomy, and regional lymph node excision. Pathology indicated adenomatous hyperplasia. The patient presented with malignant neoplasm of the thyroid. Patient history included normal delivery, alcohol abuse, and tobacco abuse. Previous surgeries included myringotomy. Patient medications included an unspecified type of birth control pills. Family history included hyperlipidemia and depressive disorder in the mother; and benign hypertension, congestive heart failure, and chronic leukemia in the grandparent(s).
UTRSNOT11	pINCY	Library was constructed using RNA isolated from uterine myometrial tissue removed from a 43-year-old female during a vaginal hysterectomy and removal of the fallopian tubes and ovaries. Pathology for the associated tumor tissue indicated that the myometrium contained an intramural and a submucosal leiomyoma. Family history included benign hypertension, hyperlipidemia, colon cancer, type II diabetes, and atherosclerotic coronary artery disease.

Table 7

Program	Description	Reference	Parameter Threshold
ABI FACTURA	A program that removes vector sequences and masks ambiguous bases in nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
ABI/PARACEL FDF	A Fast Data Finder useful in comparing and annotating amino acid or nucleic acid sequences.	Applied Biosystems, Foster City, CA; Paracel Inc., Pasadena, CA.	Mismatch <50%
ABI AutoAssembler	A program that assembles nucleic acid sequences.	Applied Biosystems, Foster City, CA.	
BLAST	A Basic Local Alignment Search Tool useful in sequence similarity search for amino acid and nucleic acid sequences. BLAST includes five functions: blastp, blastn, blastx, tblastn, and tblastx.	Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410; Altschul, S.F. et al. (1997) Nucleic Acids Res. 25:3389-3402.	*ESTs:* Probability value= 1.0E-8 or less *Full Length sequences*: Probability value= 1.0E-10 or less
FASTA	A Pearson and Lipman algorithm that searches for similarity between a query sequence and a group of sequences of the same type. FASTA comprises as least five functions: fasta, tfasta, fastx, tfastx, and ssearch.	Pearson, W.R. and D.J. Lipman (1988) Proc. Natl. Acad Sci. USA 85:2444-2448; Pearson, W.R. (1990) Methods Enzymol. 183:63-98; and Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489.	*ESTs:* fasta E value=1.06E-6 *Assembled ESTs*: fasta Identity= 95% or greater and Match length=200 bases or greater; fastx E value=1.0E-8 or less *Full Length sequences:* fastx score=100 or greater
BLIMPS	A BLocks IMProved Searcher that matches a sequence against those in BLOCKS, PRINTS, DOMO, PRODOM, and PFAM databases to search for gene families, sequence homology, and structural fingerprint regions.	Henikoff, S. and J.G. Henikoff (1991) Nucleic Acids Res. 19:6565-6572; Henikoff, J.G. and S. Henikoff (1996) Methods Enzymol. 266:88-105; and Attwood, T.K. et al. (1997) J. Chem. Inf. Comput. Sci. 37:417-424.	Probability value= 1.0E-3 or less
HMMER	An algorithm for searching a query sequence against hidden Markov model (HMM)-based databases of protein family consensus sequences, such as PFAM.	Krogh, A. et al. (1994) J. Mol. Biol. 235:1501-1531; Sonnhammer, E.L.L. et al. (1988) Nucleic Acids Res. 26:320-322; Durbin, R. et al. (1998) Our World View, in a Nutshell, Cambridge Univ. Press, pp. 1-350.	*PFAM hits:* Probability value= 1.0E-3 or less *Signal peptide hits:* Score= 0 or greater

Table 7 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
TMAP	A program that uses weight matrices to delineate transmembrane segments on protein sequences and determine orientation.	Persson, B. and P. Argos (1994) J. Mol. Biol. 237:182-192; Persson, B. and P. Argos (1996) Protein Sci. 5:363-371.	
TMHMMER	A program that uses a hidden Markov model (HMM) to delineate transmembrane segments on protein sequences and determine orientation.	Sonnhammer, E.L. et al. (1998) Proc. Sixth Intl. Conf. on Intelligent Systems for Mol. Biol., Glasgow et al., eds., The Am. Assoc. for Artificial Intelligence Press, Menlo Park, CA, pp. 175-182.	
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide selected from the group consisting of:

a) a polypeptide comprising an amino acid sequence selected from the group consisting of SEQ ID NO:1-32,

b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence selected from the group consisting of SEQ ID NO:1-32,

c) a biologically active fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, and

d) an immunogenic fragment of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1-32.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:33-64.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide selected from the group consisting of:

a) a polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64,

b) a polynucleotide comprising a naturally occurring polynucleotide sequence at least 90% identical to a polynucleotide sequence selected from the group consisting of SEQ ID NO:33-64,

c) a polynucleotide complementary to a polynucleotide of a),

d) a polynucleotide complementary to a polynucleotide of b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide has an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

18. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment the composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional TRICH, comprising administering to a patient in need of such treatment a composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1;

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, under conditions suitable for the expression of the target polynucleotide,

b) detecting altered expression of the target polynucleotide, and

c) comparing the expression of the target polynucleotide in the presence of varying amounts of the compound and in the absence of the compound.

28. A method for assessing toxicity of a test compound, said method comprising:

a) treating a biological sample containing nucleic acids with the test compound;

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

29. A diagnostic test for a condition or disease associated with the expression of TRICH in a biological sample comprising the steps of:

a) combining the biological sample with an antibody of claim 10, under conditions suitable for the antibody to bind the polypeptide and form an antibody:polypeptide complex; and

b) detecting the complex, wherein the presence of the complex correlates with the presence of the polypeptide in the biological sample.

30. The antibody of claim 10, wherein the antibody is:

a) a chimeric antibody,

b) a single chain antibody,

c) a Fab fragment,

d) a $F(ab')_2$ fragment, or

e) a humanized antibody.

31. A composition comprising an antibody of claim 10 and an acceptable excipient.

32. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 31.

33. A composition of claim 31, wherein the antibody is labeled.

34. A method of diagnosing a condition or disease associated with the expression of TRICH in a subject, comprising administering to said subject an effective amount of the composition of claim 33.

35. A method of preparing a polyclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibodies from said animal; and

c) screening the isolated antibodies with the polypeptide, thereby identifying a polyclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

36. An antibody produced by a method of claim 35.

37. A composition comprising the antibody of claim 36 and a suitable carrier.

38. A method of making a monoclonal antibody with the specificity of the antibody of claim 10 comprising:

a) immunizing an animal with a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32, or an immunogenic fragment thereof, under conditions to elicit an antibody response;

b) isolating antibody producing cells from the animal;

c) fusing the antibody producing cells with immortalized cells to form monoclonal antibody-producing hybridoma cells;

d) culturing the hybridoma cells; and

e) isolating from the culture monoclonal antibody which binds specifically to a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

39. A monoclonal antibody produced by a method of claim 38.

40. A composition comprising the antibody of claim 39 and a suitable carrier.

41. The antibody of claim 10, wherein the antibody is produced by screening a Fab expression library.

42. The antibody of claim 10, wherein the antibody is produced by screening a recombinant immunoglobulin library.

43. A method for detecting a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32 in a sample, comprising the steps of:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) detecting specific binding, wherein specific binding indicates the presence of a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32 in the sample.

44. A method of purifying a polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32 from a sample, the method comprising:

a) incubating the antibody of claim 10 with a sample under conditions to allow specific binding of the antibody and the polypeptide; and

b) separating the antibody from the sample and obtaining the purified polypeptide having an amino acid sequence selected from the group consisting of SEQ ID NO:1-32.

45. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:1.

46. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:2.

47. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:3.

48. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:4.

49. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:5.

50. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:6.

51. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:7.

52. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:8.

53. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:9.

54. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:10.

55. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:11.

56. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:12.

57. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:13.

58. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:14.

59. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:15.

60. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:16.

61. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:17.

62. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:18.

63. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:19.

64. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:20.

65. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:21.

66. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:22.

67. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:23.

68. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:24.

69. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:25.

70. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:26.

71. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:27.

72. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:28.

73. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:29.

74. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:30.

75. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:31.

76. A polypeptide of claim 1, comprising the amino acid sequence of SEQ ID NO:32.

77. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:33.

78. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:34.

79. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:35.

80. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:36.

81. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:37.

82. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:38.

83. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:39.

84. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:40.

85. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:41.

86. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:42.

87. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:43.

88. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:44.

89. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:45.

90. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:46.

91. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:47.

92. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:48.

93. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:49.

94. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:50.

95. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:51.

96. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:52.

97. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:53.

98. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:54.

99. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:55.

100. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:56.

101. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:57.

102. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:58.

103. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:59.

104. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:60.

105. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:61.

106. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:62.

107. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:63.

108. A polynucleotide of claim 11, comprising the polynucleotide sequence of SEQ ID NO:64.

<110> INCYTE GENOMICS, INC.
RAUMANN, Brigitte E.
THORNTON, Michael
DING, Li
YUE, Henry
TANG, Y.Tom
HARLAND, Lee
BURFORD, Neil
GREENE, Barrie D.
SANJANWALA, Madhu S.
BAUGHN, Mariah R.
YAO, Monique G.
YANG, Junming
PATTERSON, Chandra
GANDHI, Ameena R.
HAFALIA, April J.A.
TRIBOULEY, Catherine M.
WALIA, Narinder K.
AU-YOUNG, Janice
WALSH, Roderick T.
RAMKUMAR, Jayalaxmi
LU, Yan
LU, Dyung Aina M.
AZIMZAI, Yalda
LAL, Preeti
ELLIOTT, Vicki S.
NGUYEN, Danniel B.
XU, Yuming
SEILHAMER, Jeffrey J.
BOROWSKY, Mark L.
KHAN, Farrah A.
KEARNEY, Liam
THANGAVELU, Kavitha
DAS, Debopriya
POLICKY, Jennifer L.

<120> TRANSPORTERS AND ION CHANNELS

<130> PI-0149 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/216,547; 60/218,232; 60/220,112; 60/221,839
<151> 2000-07-07; 2000-07-14; 2000-07-21; 2000-07-28

<160> 64

<170> PERL Program

<210> 1
<211> 332
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3474673CD1

<400> 1
```
Met Tyr Arg Pro Arg Ala Arg Ala Ala Pro Glu Gly Arg Val Arg
  1               5                  10                  15
Gly Cys Ala Val Pro Ser Thr Val Leu Leu Leu Leu Ala Tyr Leu
             20                  25                  30
Ala Tyr Leu Ala Leu Gly Thr Gly Val Phe Trp Thr Leu Glu Gly
         35                  40                  45
Arg Ala Ala Gln Asp Ser Ser Arg Ser Phe Gln Arg Asp Lys Trp
```

```
                50                  55                  60
Glu Leu Leu Gln Asn Phe Thr Cys Leu Asp Arg Pro Ala Leu Asp
            65                  70                  75
Ser Leu Ile Arg Asp Val Val Gln Ala Tyr Lys Asn Gly Ala Ser
            80                  85                  90
Leu Leu Ser Asn Thr Thr Ser Met Gly Arg Trp Glu Leu Val Gly
            95                  100                 105
Ser Phe Phe Phe Ser Val Ser Thr Ile Thr Thr Ile Gly Tyr Gly
            110                 115                 120
Asn Leu Ser Pro Asn Thr Met Ala Ala Arg Leu Phe Cys Ile Phe
            125                 130                 135
Phe Ala Leu Val Gly Ile Pro Leu Asn Leu Val Val Leu Asn Arg
            140                 145                 150
Leu Gly His Leu Met Gln Gln Gly Val Asn His Trp Ala Ser Arg
            155                 160                 165
Leu Gly Gly Thr Trp Gln Asp Pro Asp Lys Ala Arg Trp Leu Ala
            170                 175                 180
Gly Ser Gly Ala Leu Leu Ser Gly Leu Leu Leu Phe Leu Leu Leu
            185                 190                 195
Pro Pro Leu Leu Phe Ser His Met Glu Gly Trp Ser Tyr Thr Glu
            200                 205                 210
Gly Phe Tyr Phe Ala Phe Ile Thr Leu Ser Thr Val Gly Phe Gly
            215                 220                 225
Asp Tyr Val Ile Gly Met Asn Pro Ser Gln Arg Tyr Pro Leu Trp
            230                 235                 240
Tyr Lys Asn Met Val Ser Leu Trp Ile Leu Phe Gly Met Ala Trp
            245                 250                 255
Leu Ala Leu Ile Ile Lys Leu Ile Leu Ser Gln Leu Glu Thr Pro
            260                 265                 270
Gly Arg Val Cys Ser Cys Cys His His Ser Ser Lys Glu Asp Phe
            275                 280                 285
Lys Ser Gln Ser Trp Arg Gln Gly Pro Asp Arg Glu Pro Glu Ser
            290                 295                 300
His Ser Pro Gln Gln Gly Cys Tyr Pro Glu Gly Pro Met Gly Ile
            305                 310                 315
Ile Gln His Leu Glu Pro Ser Ala His Ala Ala Gly Cys Gly Lys
            320                 325                 330
Asp Ser
```

<210> 2
<211> 226
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4588877CD1

<400> 2

```
Met Val Glu Met Gly Trp Asp Trp Ala Asp Arg Lys Asp Met Arg
  1                 5                  10                  15
His Arg Leu Gln Ala Gly Asn Leu Glu Asn Thr Asp Gln Val Lys
            20                  25                  30
Ser Pro Leu Leu Thr Gly Asp Ser Ser Gly Leu Pro Pro Ala Pro
            35                  40                  45
Ser Ala Pro Thr His Gly Val Lys Ala Ser Gly Gly Leu Gly Thr
            50                  55                  60
Ile Leu His Pro Gln Asp Pro Asp Lys Ala Arg Trp Leu Ala Gly
            65                  70                  75
Ser Gly Ala Leu Leu Ser Gly Leu Leu Leu Phe Leu Leu Leu Pro
            80                  85                  90
Pro Leu Leu Phe Ser His Met Glu Gly Trp Ser Tyr Thr Glu Gly
            95                  100                 105
Phe Tyr Phe Ala Phe Ile Thr Leu Ser Thr Val Gly Phe Gly Asp
            110                 115                 120
Tyr Val Ile Gly Met Asn Pro Ser Gln Arg Tyr Pro Leu Trp Tyr
```

```
                125                 130                 135
Lys Asn Met Val Ser Leu Trp Ile Leu Phe Gly Met Ala Trp Leu
                140                 145                 150
Ala Leu Ile Ile Lys Leu Ile Leu Ser Gln Leu Glu Thr Pro Gly
                155                 160                 165
Arg Val Cys Ser Cys Cys His His Ser Ser Lys Glu Asp Phe Lys
                170                 175                 180
Ser Gln Ser Trp Arg Gln Gly Pro Asp Arg Glu Pro Glu Ser His
                185                 190                 195
Ser Pro Gln Gln Gly Cys Tyr Pro Glu Gly Pro Met Gly Ile Ile
                200                 205                 210
Gln His Leu Glu Pro Ser Ala His Ala Ala Gly Cys Gly Lys Asp
                215                 220                 225
Ser
```

<210> 3
<211> 646
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472214CD1

<400> 3

```
Met Ala Glu Lys Ala Leu Glu Ala Val Gly Cys Gly Leu Gly Pro
  1                 5                  10                  15
Gly Ala Val Ala Met Ala Val Thr Leu Glu Asp Gly Ala Glu Pro
                 20                  25                  30
Pro Val Leu Thr Thr His Leu Lys Lys Val Glu Asn His Ile Thr
                 35                  40                  45
Glu Ala Gln Arg Phe Ser His Leu Pro Lys Arg Ser Ala Val Asp
                 50                  55                  60
Ile Glu Phe Val Glu Leu Ser Tyr Ser Val Arg Glu Gly Pro Cys
                 65                  70                  75
Trp Arg Lys Arg Gly Tyr Lys Thr Leu Leu Lys Cys Leu Ser Gly
                 80                  85                  90
Lys Phe Cys Arg Arg Glu Leu Ile Gly Ile Met Gly Pro Ser Gly
                 95                 100                 105
Ala Gly Lys Ser Thr Phe Met Asn Ile Leu Ala Gly Tyr Arg Glu
                110                 115                 120
Ser Gly Met Lys Gly Gln Ile Leu Val Asn Gly Arg Pro Arg Glu
                125                 130                 135
Leu Arg Thr Phe Arg Lys Met Ser Cys Tyr Ile Met Gln Asp Asp
                140                 145                 150
Met Leu Leu Pro His Leu Thr Val Leu Glu Ala Met Met Val Ser
                155                 160                 165
Ala Asn Leu Asn Leu Thr Glu Asn Pro Asp Val Lys Asn Asp Leu
                170                 175                 180
Val Thr Glu Ile Leu Thr Ala Leu Gly Leu Met Ser Cys Ser His
                185                 190                 195
Thr Arg Thr Ala Leu Leu Ser Gly Gly Gln Arg Lys Arg Leu Ala
                200                 205                 210
Ile Ala Leu Glu Leu Val Asn Asn Pro Pro Val Met Phe Phe Asp
                215                 220                 225
Glu Pro Thr Ser Gly Leu Asp Ser Ala Ser Cys Phe Gln Val Val
                230                 235                 240
Ser Leu Met Lys Ser Leu Ala Gln Gly Gly Arg Thr Ile Ile Cys
                245                 250                 255
Thr Ile His Gln Pro Ser Ala Lys Leu Phe Glu Met Phe Asp Lys
                260                 265                 270
Leu Tyr Ile Leu Ser Gln Gly Gln Cys Ile Phe Lys Gly Val Val
                275                 280                 285
Thr Asn Leu Ile Pro Tyr Leu Lys Gly Leu Gly Leu His Cys Pro
                290                 295                 300
Thr Tyr His Asn Pro Ala Asp Phe Val Ile Glu Val Ala Ser Gly
```

```
                305                 310                 315
Glu Tyr Gly Asp Leu Asn Pro Met Leu Phe Arg Ala Val Gln Asn
            320                 325                 330
Gly Leu Cys Ala Met Ala Glu Lys Lys Ser Ser Pro Glu Lys Asn
            335                 340                 345
Glu Val Pro Ala Pro Cys Pro Pro Cys Pro Pro Glu Val Asp Pro
            350                 355                 360
Ile Glu Ser His Thr Phe Ala Thr Ser Thr Leu Thr Gln Phe Cys
            365                 370                 375
Ile Leu Phe Lys Arg Thr Phe Leu Ser Ile Leu Arg Asp Thr Val
            380                 385                 390
Leu Thr His Leu Arg Phe Met Ser His Val Val Ile Gly Val Leu
            395                 400                 405
Ile Gly Leu Leu Tyr Leu His Ile Gly Asp Asp Ala Ser Lys Val
            410                 415                 420
Phe Asn Asn Thr Gly Cys Leu Phe Phe Ser Met Leu Phe Leu Met
            425                 430                 435
Phe Ala Ala Leu Met Pro Thr Val Leu Thr Val Pro Leu Glu Met
            440                 445                 450
Ala Val Phe Met Arg Glu His Leu Asn Tyr Trp Tyr Ser Leu Lys
            455                 460                 465
Ala Tyr Tyr Leu Ala Lys Thr Met Ala Asp Val Pro Phe Gln Val
            470                 475                 480
Val Cys Pro Val Val Tyr Cys Ser Ile Val Tyr Trp Met Thr Gly
            485                 490                 495
Gln Pro Ala Glu Thr Ser Arg Phe Leu Leu Phe Ser Ala Leu Ala
            500                 505                 510
Thr Ala Thr Ala Leu Val Ala Gln Ser Leu Gly Leu Leu Ile Gly
            515                 520                 525
Ala Ala Ser Asn Ser Leu Gln Val Ala Thr Phe Val Gly Pro Val
            530                 535                 540
Thr Ala Ile Pro Val Leu Leu Phe Ser Gly Phe Phe Val Ser Phe
            545                 550                 555
Lys Thr Ile Pro Thr Tyr Leu Gln Trp Ser Ser Tyr Leu Ser Tyr
            560                 565                 570
Val Arg Tyr Gly Phe Glu Gly Val Ile Leu Thr Ile Tyr Gly Met
            575                 580                 585
Glu Arg Gly Asp Leu Thr Cys Leu Glu Glu Arg Cys Pro Phe Arg
            590                 595                 600
Glu Pro Gln Ser Ile Leu Arg Ala Leu Asp Val Glu Asp Ala Lys
            605                 610                 615
Leu Tyr Met Asp Phe Leu Val Leu Gly Ile Phe Phe Leu Ala Leu
            620                 625                 630
Arg Leu Leu Ala Tyr Leu Val Leu Arg Tyr Arg Val Lys Ser Glu
            635                 640                 645
Arg
```

<210> 4
<211> 1190
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473053CD1

<400> 4

```
Met Ala Val Cys Ala Lys Lys Arg Pro Pro Glu Glu Glu Arg Arg
  1               5                  10                  15
Ala Arg Ala Asn Asp Arg Glu Tyr Asn Glu Lys Phe Gln Tyr Ala
             20                  25                  30
Ser Asn Cys Ile Lys Thr Ser Lys Tyr Asn Ile Leu Thr Phe Leu
             35                  40                  45
Pro Val Asn Leu Phe Glu Gln Phe Gln Glu Val Ala Asn Thr Tyr
             50                  55                  60
Phe Leu Phe Leu Leu Ile Leu Gln Leu Ile Pro Gln Ile Ser Ser
```

```
                65                  70                  75
Leu Ser Trp Phe Thr Thr Ile Val Pro Leu Val Leu Val Leu Thr
            80                  85                  90
Ile Thr Ala Val Lys Asp Ala Thr Asp Asp Tyr Phe Arg His Lys
        95                  100                 105
Ser Asp Asn Gln Val Asn Asn Arg Gln Ser Gln Val Leu Ile Asn
    110                 115                 120
Gly Ile Leu Gln Gln Glu Gln Trp Met Asn Val Cys Val Gly Asp
125                 130                 135
Ile Ile Lys Leu Glu Asn Asn Gln Phe Val Ala Ala Asp Leu Leu
140                 145                 150
Leu Leu Ser Ser Ser Glu Pro His Gly Leu Cys Tyr Ile Glu Thr
155                 160                 165
Ala Glu Leu Asp Gly Glu Thr Asn Met Lys Val Arg Gln Ala Ile
170                 175                 180
Pro Val Thr Ser Glu Leu Gly Asp Ile Ser Lys Leu Ala Lys Phe
185                 190                 195
Asp Gly Glu Val Ile Cys Glu Pro Pro Asn Asn Lys Leu Asp Lys
200                 205                 210
Phe Ser Gly Thr Leu Tyr Trp Lys Glu Asn Lys Phe Pro Leu Ser
215                 220                 225
Asn Gln Asn Met Leu Leu Arg Gly Cys Val Leu Arg Asn Thr Glu
230                 235                 240
Trp Cys Phe Gly Leu Val Ile Phe Ala Gly Pro Asp Thr Lys Leu
245                 250                 255
Met Gln Asn Ser Gly Arg Thr Lys Phe Lys Arg Thr Ser Ile Asp
260                 265                 270
Arg Leu Met Asn Thr Leu Val Leu Trp Ile Phe Gly Phe Leu Val
275                 280                 285
Cys Met Gly Val Ile Leu Ala Ile Gly Asn Ala Ile Trp Glu His
290                 295                 300
Glu Val Gly Met Arg Phe Gln Val Tyr Leu Pro Trp Asp Glu Ala
305                 310                 315
Val Asp Ser Ala Phe Phe Ser Gly Phe Leu Ser Phe Trp Ser Tyr
320                 325                 330
Ile Ile Ile Leu Asn Thr Val Val Pro Ile Ser Leu Tyr Val Ser
335                 340                 345
Val Glu Val Ile Arg Leu Gly His Ser Tyr Phe Ile Asn Trp Asp
350                 355                 360
Lys Lys Met Phe Cys Met Lys Lys Arg Thr Pro Ala Glu Ala Arg
365                 370                 375
Thr Thr Thr Leu Asn Glu Glu Leu Gly Gln Val Glu Tyr Ile Phe
380                 385                 390
Ser Asp Lys Thr Gly Thr Leu Thr Gln Asn Ile Met Val Phe Asn
395                 400                 405
Lys Cys Ser Ile Asn Gly His Ser Tyr Gly Asp Val Phe Asp Val
410                 415                 420
Leu Gly His Lys Ala Glu Leu Gly Glu Arg Pro Glu Pro Val Asp
425                 430                 435
Phe Ser Phe Asn Pro Leu Ala Asp Lys Lys Phe Leu Phe Trp Asp
440                 445                 450
Pro Ser Leu Leu Glu Ala Val Lys Ile Gly Asp Pro His Thr His
455                 460                 465
Glu Phe Phe Arg Leu Leu Ser Leu Cys His Thr Val Met Ser Glu
470                 475                 480
Glu Lys Asn Glu Gly Glu Leu Tyr Tyr Lys Ala Gln Ser Pro Asp
485                 490                 495
Glu Gly Ala Leu Val Thr Ala Ala Arg Asn Phe Gly Phe Val Phe
500                 505                 510
Arg Ser Arg Thr Pro Lys Thr Ile Thr Val His Glu Met Gly Thr
515                 520                 525
Ala Ile Thr Tyr Gln Leu Leu Ala Ile Leu Asp Phe Asn Asn Ile
530                 535                 540
Arg Lys Arg Met Ser Val Ile Val Arg Asn Pro Glu Gly Lys Ile
545                 550                 555
Arg Leu Tyr Cys Lys Gly Ala Asp Thr Ile Leu Leu Asp Arg Leu
560                 565                 570
```

```
His His Ser Thr Gln Glu Leu Leu Asn Thr Thr Met Asp His Leu
575                 580                 585
Asn Glu Tyr Ala Gly Glu Gly Leu Arg Thr Leu Val Leu Ala Tyr
590                 595                 600
Lys Asp Leu Asp Glu Glu Tyr Tyr Glu Glu Trp Ala Glu Arg Arg
605                 610                 615
Leu Gln Ala Ser Leu Ala Gln Asp Ser Arg Glu Asp Arg Leu Ala
620                 625                 630
Ser Ile Tyr Glu Glu Val Glu Asn Asn Met Met Leu Leu Gly Ala
635                 640                 645
Thr Ala Ile Glu Asp Lys Leu Gln Gln Gly Val Pro Glu Thr Ile
650                 655                 660
Ala Leu Leu Thr Leu Ala Asn Ile Lys Ile Trp Val Leu Thr Gly
665                 670                 675
Asp Lys Gln Glu Thr Ala Val Asn Ile Gly Tyr Ser Cys Lys Met
680                 685                 690
Leu Thr Asp Asp Met Thr Glu Val Phe Ile Val Thr Gly His Thr
695                 700                 705
Val Leu Glu Val Arg Glu Glu Leu Arg Lys Ala Arg Glu Lys Met
710                 715                 720
Met Asp Ser Ser Arg Ser Val Gly Asn Gly Phe Thr Tyr Gln Asp
725                 730                 735
Lys Leu Ser Ser Ser Lys Leu Thr Ser Val Leu Glu Ala Val Ala
740                 745                 750
Gly Glu Tyr Ala Leu Val Ile Asn Gly His Ser Leu Ala His Ala
755                 760                 765
Leu Glu Ala Asp Met Glu Leu Glu Phe Leu Glu Thr Ala Cys Ala
770                 775                 780
Cys Lys Ala Val Ile Cys Cys Arg Val Thr Pro Leu Gln Lys Ala
785                 790                 795
Gln Val Val Glu Leu Val Lys Lys Tyr Lys Lys Ala Val Thr Leu
800                 805                 810
Ala Ile Gly Asp Gly Ala Asn Asp Val Ser Met Ile Lys Thr Ala
815                 820                 825
His Ile Gly Val Gly Ile Ser Gly Gln Glu Gly Ile Gln Ala Val
830                 835                 840
Leu Ala Ser Asp Tyr Ser Phe Ser Gln Phe Lys Phe Leu Gln Arg
845                 850                 855
Leu Leu Leu Val His Gly Arg Trp Ser Tyr Leu Arg Met Cys Lys
860                 865                 870
Phe Leu Cys Tyr Phe Phe Tyr Lys Asn Phe Ala Phe Thr Met Val
875                 880                 885
His Phe Trp Phe Gly Phe Phe Cys Gly Phe Ser Ala Gln Thr Val
890                 895                 900
Tyr Asp Gln Tyr Phe Ile Thr Leu Tyr Asn Ile Val Tyr Thr Ser
905                 910                 915
Leu Pro Val Leu Ala Met Gly Val Phe Asp Gln Asp Val Pro Glu
920                 925                 930
Gln Arg Ser Met Glu Tyr Pro Lys Leu Tyr Glu Pro Gly Gln Leu
935                 940                 945
Asn Leu Leu Phe Asn Lys Arg Glu Phe Phe Ile Cys Ile Ala Gln
950                 955                 960
Gly Ile Tyr Thr Ser Val Leu Met Phe Phe Ile Pro Tyr Gly Val
965                 970                 975
Phe Ala Asp Ala Thr Arg Asp Asp Gly Thr Gln Leu Ala Asp Tyr
980                 985                 990
Gln Ser Phe Ala Val Thr Val Ala Thr Ser Leu Val Ile Val Val
995                1000                1005
Ser Val Gln Ile Gly Leu Asp Thr Gly Tyr Trp Thr Ala Ile Asn
1010                1015                1020
His Phe Phe Ile Trp Gly Ser Leu Ala Val Tyr Phe Ala Ile Leu
1025                1030                1035
Phe Ala Met His Ser Asn Gly Leu Phe Asp Met Phe Pro Asn Gln
1040                1045                1050
Phe Arg Phe Val Gly Asn Ala Gln Asn Thr Leu Ala Gln Pro Thr
1055                1060                1065
Val Trp Leu Thr Ile Val Leu Thr Thr Val Val Cys Ile Met Pro
```

```
                1070                1075                1080
Val Val Ala Phe Arg Phe Leu Arg Leu Asn Leu Lys Pro Asp Leu
                1085                1090                1095
Ser Asp Thr Val Arg Tyr Thr Gln Leu Val Arg Lys Lys Gln Lys
                1100                1105                1110
Ala Gln His Arg Cys Met Arg Arg Val Gly Arg Thr Gly Ser Arg
                1115                1120                1125
Arg Ser Gly Tyr Ala Phe Ser His Gln Glu Gly Phe Gly Glu Leu
                1130                1135                1140
Ile Met Ser Gly Lys Asn Met Arg Leu Ser Ser Leu Ala Leu Ser
                1145                1150                1155
Ser Phe Thr Thr Arg Ser Ser Ser Ser Trp Ile Glu Ser Leu Arg
                1160                1165                1170
Arg Lys Lys Ser Asp Ser Ala Ser Ser Pro Ser Gly Gly Ala Asp
                1175                1180                1185
Lys Pro Leu Lys Gly
                1190
```

<210> 5
<211> 467
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473347CD1

<400> 5

```
Met Val Leu Ala Phe Gln Leu Val Ser Phe Thr Tyr Ile Trp Ile
  1                 5                  10                  15
Ile Leu Lys Pro Asn Val Cys Ala Ala Ser Asn Ile Lys Met Thr
                 20                  25                  30
His Gln Arg Cys Ser Ser Ser Met Lys Gln Thr Cys Lys Gln Glu
                 35                  40                  45
Thr Arg Met Lys Lys Asp Asp Ser Thr Lys Ala Arg Pro Gln Lys
                 50                  55                  60
Tyr Glu Gln Leu Leu His Ile Glu Asp Asn Asp Phe Ala Met Arg
                 65                  70                  75
Pro Gly Phe Gly Gly Ser Pro Val Pro Val Gly Ile Asp Val His
                 80                  85                  90
Val Glu Ser Ile Asp Ser Ile Ser Glu Thr Asn Met Asp Phe Thr
                 95                 100                 105
Met Thr Phe Tyr Leu Arg His Tyr Trp Lys Asp Glu Arg Leu Ser
                110                 115                 120
Phe Pro Ser Thr Ala Asn Lys Ser Met Thr Phe Asp His Arg Leu
                125                 130                 135
Thr Arg Lys Ile Trp Val Pro Asp Ile Phe Phe Val His Ser Lys
                140                 145                 150
Arg Ser Phe Ile His Asp Thr Thr Met Glu Asn Ile Met Leu Arg
                155                 160                 165
Val His Pro Asp Gly Asn Val Leu Leu Ser Leu Arg Ile Thr Val
                170                 175                 180
Ser Ala Met Cys Phe Met Asp Phe Ser Arg Phe Pro Leu Asp Thr
                185                 190                 195
Gln Asn Cys Ser Leu Glu Leu Glu Ser Tyr Ala Tyr Asn Glu Asp
                200                 205                 210
Asp Leu Met Leu Tyr Trp Lys His Gly Asn Lys Ser Leu Asn Thr
                215                 220                 225
Glu Glu His Met Ser Leu Ser Gln Phe Phe Ile Glu Asp Phe Ser
                230                 235                 240
Ala Ser Ser Gly Leu Ala Phe Tyr Ser Ser Thr Gly Trp Tyr Asn
                245                 250                 255
Arg Leu Phe Ile Ile Ser Val Leu Arg Arg His Val Phe Phe Phe
                260                 265                 270
Val Leu Pro Thr Tyr Tyr Pro Ala Ile Leu Met Val Met Leu Ser
                275                 280                 285
Trp Val Ser Phe Trp Ile Asp Arg Arg Ala Val Pro Ala Arg Val
```

```
                290                 295                 300
Ser Leu Gly Ile Thr Thr Val Leu Thr Met Ser Thr Ile Ile Thr
                305                 310                 315
Ala Val Ser Ala Ser Met Pro Gln Val Ser Tyr Leu Lys Ala Val
                320                 325                 330
Asp Val Tyr Leu Trp Val Ser Ser Leu Phe Val Phe Leu Ser Val
                335                 340                 345
Ile Glu Tyr Ala Ala Val Asn Tyr Leu Thr Thr Val Glu Glu Arg
                350                 355                 360
Lys Gln Phe Lys Lys Thr Gly Lys Ile Ser Arg Met Tyr Asn Ile
                365                 370                 375
Asp Ala Val Gln Ala Met Ala Phe Asp Gly Cys Tyr His Asp Ser
                380                 385                 390
Glu Ile Asp Met Asp Gln Thr Ser Leu Ser Leu Asn Ser Glu Asp
                395                 400                 405
Phe Met Arg Arg Lys Ser Ile Cys Ser Pro Ser Thr Asp Ser Ser
                410                 415                 420
Arg Ile Lys Arg Arg Lys Ser Leu Gly Gly His Val Gly Arg Ile
                425                 430                 435
Ile Leu Glu Asn Asn His Val Ile Asp Thr Tyr Ser Arg Ile Leu
                440                 445                 450
Phe Pro Ile Val Tyr Ile Leu Phe Asn Leu Phe Tyr Trp Gly Val
                455                 460                 465
Tyr Val
```

<210> 6
<211> 1196
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474240CD1

<400> 6

```
Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Phe Leu
  1                 5                  10                  15
Gly Thr Ile Ile Arg Lys Phe Glu Gly Gln Asn Lys Lys Phe Ile
                 20                  25                  30
Ile Ala Asn Ala Arg Val Gln Asn Cys Ala Ile Ile Tyr Cys Asn
                 35                  40                  45
Asp Gly Phe Cys Glu Met Thr Gly Phe Ser Arg Pro Asp Val Met
                 50                  55                  60
Gln Lys Pro Cys Thr Cys Asp Phe Leu His Gly Pro Glu Thr Lys
                 65                  70                  75
Arg His Asp Ile Ala Gln Ile Ala Gln Ala Leu Leu Gly Ser Glu
                 80                  85                  90
Glu Arg Lys Val Glu Val Thr Tyr Tyr His Lys Asn Gly Ser Thr
                 95                 100                 105
Phe Ile Cys Asn Thr His Ile Ile Pro Val Lys Asn Gln Glu Gly
                110                 115                 120
Val Ala Met Met Phe Ile Ile Asn Phe Glu Tyr Val Thr Asp Asn
                125                 130                 135
Glu Asn Ala Ala Thr Pro Glu Arg Val Asn Pro Ile Leu Pro Ile
                140                 145                 150
Lys Thr Val Asn Arg Lys Phe Phe Gly Phe Lys Phe Pro Gly Leu
                155                 160                 165
Arg Val Leu Thr Tyr Arg Lys Gln Ser Leu Pro Gln Glu Asp Pro
                170                 175                 180
Asp Val Val Val Ile Asp Ser Ser Lys His Ser Asp Asp Ser Val
                185                 190                 195
Ala Met Lys His Phe Lys Ser Pro Thr Lys Glu Ser Cys Ser Pro
                200                 205                 210
Ser Glu Ala Asp Asp Thr Lys Ala Leu Ile Gln Pro Ser Lys Cys
                215                 220                 225
Ser Pro Leu Val Asn Ile Ser Gly Pro Leu Asp His Ser Ser Pro
```

```
                230                 235                 240
Lys Arg Gln Trp Asp Arg Leu Tyr Pro Asp Met Leu Gln Ser Ser
                245                 250                 255
Ser Gln Leu Ser His Ser Arg Ser Arg Glu Ser Leu Cys Ser Ile
                260                 265                 270
Arg Arg Ala Ser Ser Val His Asp Ile Glu Gly Phe Gly Val His
                275                 280                 285
Pro Lys Asn Ile Phe Arg Asp Arg His Ala Ser Glu Asp Asn Gly
                290                 295                 300
Arg Asn Val Lys Gly Pro Phe Asn His Ile Lys Ser Ser Leu Leu
                305                 310                 315
Gly Ser Thr Ser Asp Ser Asn Leu Asn Lys Tyr Ser Thr Ile Asn
                320                 325                 330
Lys Ile Pro Gln Leu Thr Leu Asn Phe Ser Glu Val Lys Thr Glu
                335                 340                 345
Lys Lys Asn Ser Ser Pro Pro Ser Ser Asp Lys Thr Ile Ile Ala
                350                 355                 360
Pro Lys Val Lys Asp Arg Thr His Asn Val Thr Glu Lys Val Thr
                365                 370                 375
Gln Val Leu Ser Leu Gly Ala Asp Val Leu Pro Glu Tyr Lys Leu
                380                 385                 390
Gln Thr Pro Arg Ile Asn Lys Phe Thr Ile Leu His Tyr Ser Pro
                395                 400                 405
Phe Lys Ala Val Trp Asp Trp Leu Ile Leu Leu Leu Val Ile Tyr
                410                 415                 420
Thr Ala Ile Phe Thr Pro Tyr Ser Ala Ala Phe Leu Leu Asn Asp
                425                 430                 435
Arg Glu Glu Gln Lys Arg Arg Glu Cys Gly Tyr Ser Cys Ser Pro
                440                 445                 450
Leu Asn Val Val Asp Leu Ile Val Asp Ile Met Phe Ile Ile Asp
                455                 460                 465
Ile Leu Ile Asn Phe Arg Thr Thr Tyr Val Asn Gln Asn Glu Glu
                470                 475                 480
Val Val Ser Asp Pro Ala Lys Ile Ala Ile His Tyr Phe Lys Gly
                485                 490                 495
Trp Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu Leu
                500                 505                 510
Ile Phe Gly Ser Gly Ser Asp Glu Thr Thr Thr Leu Ile Gly Leu
                515                 520                 525
Leu Lys Thr Ala Arg Leu Leu Arg Leu Val Arg Val Ala Arg Lys
                530                 535                 540
Leu Asp Arg Tyr Ser Glu Tyr Gly Ala Ala Val Leu Met Leu Leu
                545                 550                 555
Met Cys Ile Phe Ala Leu Ile Ala His Trp Leu Ala Cys Ile Trp
                560                 565                 570
Tyr Ala Ile Gly Asn Val Glu Arg Pro Tyr Leu Thr Asp Lys Ile
                575                 580                 585
Gly Trp Leu Asp Ser Leu Gly Gln Gln Ile Gly Lys Arg Tyr Asn
                590                 595                 600
Asp Ser Asp Ser Ser Ser Gly Pro Ser Ile Lys Asp Lys Tyr Val
                605                 610                 615
Thr Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr Ser Val Gly Phe
                620                 625                 630
Gly Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Ile Phe Ser Ile
                635                 640                 645
Cys Val Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile Phe Gly
                650                 655                 660
Asn Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala Arg
                665                 670                 675
Tyr His Met Gln Met Leu Arg Val Lys Glu Phe Ile Arg Phe His
                680                 685                 690
Gln Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln
                695                 700                 705
His Ala Trp Thr Tyr Thr Asn Gly Ile Asp Met Asn Met Val Leu
                710                 715                 720
Lys Gly Phe Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu
                725                 730                 735
```

```
Asn Gln Thr Leu Leu Gln Asn Cys Lys Ala Phe Arg Gly Ala Ser
                740                 745                 750
Lys Gly Cys Leu Arg Ala Leu Ala Met Lys Phe Lys Thr Thr His
                755                 760                 765
Ala Pro Pro Gly Asp Thr Leu Val His Cys Gly Asp Val Leu Thr
                770                 775                 780
Ala Leu Tyr Phe Leu Ser Arg Gly Ser Ile Glu Ile Leu Lys Asp
                785                 790                 795
Asp Ile Val Val Ala Ile Leu Gly Lys Asn Asp Ile Phe Gly Glu
                800                 805                 810
Met Val His Leu Tyr Ala Lys Pro Gly Lys Ser Asn Ala Asp Val
                815                 820                 825
Arg Ala Leu Thr Tyr Cys Asp Leu His Lys Ile Gln Arg Glu Asp
                830                 835                 840
Leu Leu Glu Val Leu Asp Met Tyr Pro Glu Phe Ser Asp His Phe
                845                 850                 855
Leu Thr Asn Leu Glu Leu Thr Phe Asn Leu Arg His Glu Ser Ala
                860                 865                 870
Lys Ala Asp Leu Leu Arg Ser Gln Ser Met Asn Asp Ser Glu Gly
                875                 880                 885
Asp Asn Cys Lys Leu Arg Arg Arg Lys Leu Ser Phe Glu Ser Glu
                890                 895                 900
Gly Glu Lys Glu Asn Ser Thr Asn Asp Pro Glu Asp Ser Ala Asp
                905                 910                 915
Thr Ile Arg His Tyr Gln Ser Ser Lys Arg His Phe Glu Glu Lys
                920                 925                 930
Lys Ser Arg Ser Ser Ser Phe Ile Ser Ser Ile Asp Asp Glu Gln
                935                 940                 945
Lys Pro Leu Phe Ser Gly Ile Val Asp Ser Ser Pro Gly Ile Gly
                950                 955                 960
Lys Ala Ser Gly Leu Asp Phe Glu Glu Thr Val Pro Thr Ser Gly
                965                 970                 975
Arg Met His Ile Asp Lys Arg Ser His Ser Cys Lys Asp Ile Thr
                980                 985                 990
Asp Met Arg Ser Trp Glu Arg Glu Asn Ala His Pro Gln Pro Glu
                995                1000                1005
Asp Ser Ser Pro Ser Ala Leu Gln Arg Ala Ala Trp Gly Ile Ser
               1010                1015                1020
Glu Thr Glu Ser Asp Leu Thr Tyr Gly Glu Val Glu Gln Arg Leu
               1025                1030                1035
Asp Leu Leu Gln Glu Gln Leu Asn Arg Leu Glu Ser Gln Met Thr
               1040                1045                1050
Thr Asp Ile Gln Thr Ile Leu Gln Leu Leu Gln Lys Gln Thr Thr
               1055                1060                1065
Val Val Pro Pro Ala Tyr Ser Met Val Thr Ala Gly Ser Glu Tyr
               1070                1075                1080
Gln Arg Pro Ile Ile Gln Leu Met Arg Thr Ser Gln Pro Glu Ala
               1085                1090                1095
Ser Ile Lys Thr Asp Arg Ser Phe Ser Pro Ser Ser Gln Cys Pro
               1100                1105                1110
Glu Phe Leu Asp Leu Glu Lys Ser Lys Leu Lys Ser Lys Glu Ser
               1115                1120                1125
Leu Ser Ser Gly Val His Leu Asn Thr Ala Ser Glu Asp Asn Leu
               1130                1135                1140
Thr Ser Leu Leu Lys Gln Asp Ser Asp Leu Ser Leu Glu Leu His
               1145                1150                1155
Leu Arg Gln Arg Lys Thr Tyr Val His Pro Ile Arg His Pro Ser
               1160                1165                1170
Leu Pro Asp Ser Ser Leu Ser Thr Val Gly Ile Val Gly Leu His
               1175                1180                1185
Arg His Val Ser Asp Pro Gly Leu Pro Gly Lys
               1190                1195
```

<210> 7
<211> 512
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7475338CD1

<400> 7

```
Met Glu Asn Lys Glu Ala Gly Thr Pro Pro Pro Ile Pro Ser Arg
  1               5                  10                  15
Glu Gly Arg Leu Gln Pro Thr Leu Leu Leu Ala Thr Leu Ser Ala
             20                  25                  30
Ala Phe Gly Ser Ala Phe Gln Tyr Gly Tyr Asn Leu Ser Val Val
         35                  40                  45
Asn Thr Pro His Lys Val Phe Lys Ser Phe Tyr Asn Glu Thr Tyr
     50                  55                  60
Phe Glu Arg His Ala Thr Phe Met Asp Gly Lys Leu Met Leu Leu
 65                  70                  75
Leu Trp Ser Cys Thr Val Ser Met Phe Pro Leu Gly Gly Leu Leu
             80                  85                  90
Gly Ser Leu Leu Val Gly Leu Leu Val Asp Ser Cys Gly Arg Lys
         95                 100                 105
Gly Thr Leu Leu Ile Asn Asn Ile Phe Ala Ile Ile Pro Ala Ile
    110                 115                 120
Leu Met Gly Val Ser Lys Val Ala Lys Ala Phe Glu Leu Ile Val
125                 130                 135
Phe Ser Arg Val Val Leu Gly Val Cys Ala Gly Ile Ser Tyr Ser
            140                 145                 150
Ala Leu Pro Met Tyr Leu Gly Glu Leu Ala Pro Lys Asn Leu Arg
        155                 160                 165
Gly Met Val Gly Thr Met Thr Glu Val Phe Val Ile Val Gly Val
    170                 175                 180
Phe Leu Ala Gln Ile Phe Ser Leu Gln Ala Ile Leu Gly Asn Pro
185                 190                 195
Ala Gly Trp Pro Val Leu Leu Ala Leu Thr Gly Val Pro Ala Leu
            200                 205                 210
Leu Gln Leu Leu Thr Leu Pro Phe Phe Pro Glu Ser Pro Arg Tyr
        215                 220                 225
Ser Leu Ile Gln Lys Gly Asp Glu Ala Thr Ala Arg Gln Ala Leu
    230                 235                 240
Arg Arg Leu Arg Gly His Thr Asp Met Glu Ala Glu Leu Glu Asp
245                 250                 255
Met Arg Ala Glu Ala Arg Ala Glu Arg Ala Glu Gly His Leu Ser
            260                 265                 270
Val Leu His Leu Cys Ala Leu Arg Ser Leu Arg Trp Gln Leu Leu
        275                 280                 285
Ser Ile Ile Val Leu Met Ala Gly Gln Gln Leu Ser Gly Ile Asn
    290                 295                 300
Ala Ile Asn Tyr Tyr Ala Asp Thr Ile Tyr Thr Ser Ala Gly Val
305                 310                 315
Glu Ala Ala His Ser Gln Tyr Val Thr Val Gly Ser Gly Val Val
            320                 325                 330
Asn Ile Val Met Thr Ile Thr Ser Ala Val Leu Val Glu Arg Leu
        335                 340                 345
Gly Arg Arg His Leu Leu Leu Ala Gly Tyr Gly Ile Cys Gly Ser
    350                 355                 360
Ala Cys Leu Val Leu Thr Val Val Leu Leu Phe Gln Asn Arg Val
365                 370                 375
Pro Glu Leu Ser Tyr Leu Gly Ile Ile Cys Val Phe Ala Tyr Ile
            380                 385                 390
Ala Gly His Ser Ile Gly Pro Ser Pro Val Pro Ser Val Val Arg
        395                 400                 405
Thr Glu Ile Phe Leu Gln Ser Ser Arg Arg Ala Ala Phe Met Val
    410                 415                 420
Asp Gly Ala Val His Trp Leu Thr Asn Phe Ile Ile Gly Phe Leu
425                 430                 435
Phe Pro Ser Ile Gln Glu Ala Ile Gly Ala Tyr Ser Phe Ile Ile
            440                 445                 450
Phe Ala Gly Ile Cys Leu Leu Thr Ala Ile Tyr Ile Tyr Val Val
        455                 460                 465
```

```
Ile Pro Glu Thr Lys Gly Lys Thr Phe Val Glu Ile Asn Arg Ile
                470                 475                 480
Phe Ala Lys Arg Asn Arg Val Lys Leu Pro Glu Glu Lys Glu Glu
                485                 490                 495
Thr Ile Asp Ala Gly Pro Pro Thr Ala Ser Pro Ala Lys Glu Thr
                500                 505                 510
Ser Phe
```

<210> 8
<211> 568
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476747CD1

<400> 8

```
Met Thr Ala Ser Thr Pro Glu Ala Thr Pro Asn Met Glu Leu Lys
  1               5                  10                  15
Ala Pro Ala Ala Gly Gly Leu Asn Ala Gly Pro Val Pro Pro Ala
                20                  25                  30
Ala Met Ser Thr Gln Arg Leu Arg Asn Glu Asp Tyr His Asp Tyr
                35                  40                  45
Ser Ser Thr Asp Val Ser Pro Glu Glu Ser Pro Ser Glu Gly Leu
                50                  55                  60
Asn Asn Leu Ser Ser Pro Gly Ser Tyr Gln Arg Phe Gly Gln Ser
                65                  70                  75
Asn Ser Thr Thr Trp Phe Gln Thr Leu Ile His Leu Leu Lys Gly
                80                  85                  90
Asn Ile Gly Thr Gly Leu Leu Gly Leu Pro Leu Ala Val Lys Asn
                95                 100                 105
Ala Gly Ile Val Met Gly Pro Ile Ser Leu Leu Ile Ile Gly Ile
               110                 115                 120
Val Ala Val His Cys Met Gly Ile Leu Val Lys Cys Ala His His
               125                 130                 135
Phe Cys Arg Arg Leu Asn Lys Ser Phe Val Asp Tyr Gly Asp Thr
               140                 145                 150
Val Met Tyr Gly Leu Glu Ser Ser Pro Cys Ser Trp Leu Arg Asn
               155                 160                 165
His Ala His Trp Gly Arg Arg Val Val Asp Phe Phe Leu Ile Val
               170                 175                 180
Thr Gln Leu Gly Phe Cys Cys Val Tyr Phe Val Phe Leu Ala Asp
               185                 190                 195
Asn Phe Lys Gln Val Ile Glu Ala Ala Asn Gly Thr Thr Asn Asn
               200                 205                 210
Cys His Asn Asn Glu Thr Val Ile Leu Thr Pro Thr Met Asp Ser
               215                 220                 225
Arg Leu Tyr Met Leu Ser Phe Leu Pro Phe Leu Val Leu Leu Val
               230                 235                 240
Phe Ile Arg Asn Leu Arg Ala Leu Ser Ile Phe Ser Leu Leu Ala
               245                 250                 255
Asn Ile Thr Met Leu Val Ser Leu Val Met Ile Tyr Gln Phe Ile
               260                 265                 270
Val Gln Arg Ile Pro Asp Pro Ser His Leu Pro Leu Val Ala Pro
               275                 280                 285
Trp Lys Thr Tyr Pro Leu Phe Phe Gly Thr Ala Ile Phe Ser Phe
               290                 295                 300
Glu Gly Ile Gly Met Val Leu Pro Leu Glu Asn Lys Met Lys Asp
               305                 310                 315
Pro Arg Lys Phe Pro Leu Ile Leu Tyr Leu Gly Met Val Ile Val
               320                 325                 330
Thr Ile Leu Tyr Ile Ser Leu Gly Cys Leu Gly Tyr Leu Gln Phe
               335                 340                 345
Gly Ala Asn Ile Gln Gly Ser Ile Thr Leu Asn Leu Pro Asn Cys
               350                 355                 360
```

```
Trp Leu Tyr Gln Ser Val Lys Leu Leu Tyr Ser Ile Gly Ile Phe
                365                 370                 375
Phe Thr Tyr Ala Leu Gln Phe Tyr Val Pro Ala Glu Ile Ile Ile
                380                 385                 390
Pro Phe Phe Val Ser Arg Ala Pro Glu Pro Cys Glu Leu Val Val
                395                 400                 405
Asp Leu Phe Val Arg Pro Val Leu Val Cys Leu Thr Ser Leu Ser
                410                 415                 420
Gly Ser Val Asp Asn Gly Trp Tyr Gly Thr Glu Ala Asp Gly Thr
                425                 430                 435
Ser Cys Gly Ser Ala Pro Leu Val Phe Val Ser Ser Ser Phe Leu
                440                 445                 450
Ala His Pro Trp Leu Ser Phe Arg Cys Glu Ser Gln Trp Val Ser
                455                 460                 465
Cys His Arg Asp Thr Val Val Val Trp Gly Phe Ala Arg Gly Ile
                470                 475                 480
Leu Ala Ile Leu Ile Pro Arg Leu Asp Leu Val Ile Ser Leu Val
                485                 490                 495
Gly Ser Val Ser Ser Ser Ala Leu Ala Leu Ile Ile Pro Pro Leu
                500                 505                 510
Leu Glu Val Thr Thr Phe Tyr Ser Glu Gly Met Ser Pro Leu Thr
                515                 520                 525
Ile Phe Lys Asp Ala Leu Ile Ser Ile Leu Gly Phe Val Gly Phe
                530                 535                 540
Val Val Gly Thr Tyr Glu Ala Leu Tyr Glu Leu Ile Gln Pro Ser
                545                 550                 555
Asn Ala Pro Ile Phe Ile Asn Ser Thr Cys Ala Phe Ile
                560                 565

<210> 9
<211> 958
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477898CD1

<400> 9
Met Pro Val Arg Arg Gly His Val Ala Pro Gln Asn Thr Tyr Leu
  1                 5                  10                  15
Asp Thr Ile Ile Arg Lys Phe Glu Gly Gln Ser Arg Lys Phe Leu
                 20                  25                  30
Ile Ala Asn Ala Gln Met Glu Asn Cys Ala Ile Ile Tyr Cys Asn
                 35                  40                  45
Asp Gly Phe Cys Glu Leu Phe Gly Tyr Ser Arg Val Glu Val Met
                 50                  55                  60
Gln Gln Pro Cys Thr Cys Asp Phe Leu Thr Gly Pro Asn Thr Pro
                 65                  70                  75
Ser Ser Ala Val Ser Arg Leu Ala Gln Ala Leu Leu Gly Ala Glu
                 80                  85                  90
Glu Cys Lys Val Asp Ile Leu Tyr Tyr Arg Lys Asp Ala Ser Ser
                 95                 100                 105
Phe Arg Cys Leu Val Asp Val Val Pro Val Lys Asn Glu Asp Gly
                110                 115                 120
Ala Val Ile Met Phe Ile Leu Asn Phe Glu Asp Leu Ala Gln Leu
                125                 130                 135
Leu Ala Lys Cys Ser Ser Arg Ser Leu Ser Gln Arg Leu Leu Ser
                140                 145                 150
Gln Ser Phe Leu Gly Ser Glu Gly Ser His Gly Arg Pro Gly Gly
                155                 160                 165
Pro Gly Pro Gly Thr Gly Arg Gly Lys Tyr Arg Thr Ile Ser Gln
                170                 175                 180
Ile Pro Gln Phe Thr Leu Asn Phe Val Glu Phe Asn Leu Glu Lys
                185                 190                 195
His Arg Ser Ser Ser Thr Thr Glu Ile Glu Ile Ile Ala Pro His
                200                 205                 210
```

```
Lys Val Val Glu Arg Thr Gln Asn Val Thr Glu Lys Val Thr Gln
215                 220                 225
Val Leu Ser Leu Gly Ala Asp Val Leu Pro Glu Tyr Lys Leu Gln
230                 235                 240
Ala Pro Arg Ile His Arg Trp Thr Ile Leu His Tyr Ser Pro Phe
245                 250                 255
Lys Ala Val Trp Asp Trp Leu Ile Leu Leu Leu Val Ile Tyr Thr
260                 265                 270
Ala Val Phe Thr Pro Tyr Ser Ala Ala Phe Leu Leu Ser Asp Gln
275                 280                 285
Asp Glu Ser Arg Arg Gly Ala Cys Ser Tyr Thr Cys Ser Pro Leu
290                 295                 300
Thr Val Val Asp Leu Ile Val Asp Ile Met Phe Val Val Asp Ile
305                 310                 315
Val Ile Asn Phe Arg Thr Thr Tyr Val Asn Thr Asn Asp Glu Val
320                 325                 330
Val Ser His Pro Arg Arg Ile Ala Val His Tyr Phe Lys Gly Trp
335                 340                 345
Phe Leu Ile Asp Met Val Ala Ala Ile Pro Phe Asp Leu Leu Ile
350                 355                 360
Phe Arg Thr Gly Ser Asp Glu Thr Thr Thr Leu Ile Gly Leu Leu
365                 370                 375
Lys Thr Ala Arg Leu Leu Arg Leu Val Arg Val Ala Arg Lys Leu
380                 385                 390
Asp Arg Tyr Ser Glu Tyr Gly Ala Ala Val Leu Phe Leu Leu Met
395                 400                 405
Cys Thr Phe Pro Leu Ile Ala His Trp Leu Ala Cys Ile Trp Tyr
410                 415                 420
Ala Ile Gly Asn Val Glu Arg Pro Tyr Leu Glu His Lys Ile Gly
425                 430                 435
Trp Leu Asp Ser Leu Gly Val Gln Leu Gly Lys Arg Tyr Asn Gly
440                 445                 450
Ser Asp Pro Ala Ser Gly Pro Ser Val Gln Asp Lys Tyr Val Thr
455                 460                 465
Ala Leu Tyr Phe Thr Phe Ser Ser Leu Thr Ser Val Gly Phe Gly
470                 475                 480
Asn Val Ser Pro Asn Thr Asn Ser Glu Lys Val Phe Ser Ile Cys
485                 490                 495
Val Met Leu Ile Gly Ser Leu Met Tyr Ala Ser Ile Phe Gly Asn
500                 505                 510
Val Ser Ala Ile Ile Gln Arg Leu Tyr Ser Gly Thr Ala Arg Tyr
515                 520                 525
His Thr Gln Met Leu Arg Val Lys Glu Phe Ile Arg Phe His Gln
530                 535                 540
Ile Pro Asn Pro Leu Arg Gln Arg Leu Glu Glu Tyr Phe Gln His
545                 550                 555
Ala Trp Ser Tyr Thr Asn Gly Ile Asp Met Asn Ala Val Leu Lys
560                 565                 570
Gly Phe Pro Glu Cys Leu Gln Ala Asp Ile Cys Leu His Leu His
575                 580                 585
Arg Ala Leu Leu Gln His Cys Pro Ala Phe Ser Gly Ala Gly Lys
590                 595                 600
Gly Cys Leu Arg Ala Leu Ala Val Lys Phe Lys Thr Thr His Ala
605                 610                 615
Pro Pro Gly Asp Thr Leu Val His Leu Gly Asp Val Leu Ser Thr
620                 625                 630
Leu Tyr Phe Ile Ser Arg Gly Ser Ile Glu Ile Leu Arg Asp Asp
635                 640                 645
Val Val Val Ala Ile Leu Gly Lys Asn Asp Ile Phe Gly Glu Pro
650                 655                 660
Val Ser Leu His Ala Gln Pro Gly Lys Ser Ser Ala Asp Val Arg
665                 670                 675
Ala Leu Thr Tyr Cys Asp Leu His Lys Ile Gln Arg Ala Asp Leu
680                 685                 690
Leu Glu Val Leu Asp Met Tyr Pro Ala Phe Ala Glu Ser Phe Trp
695                 700                 705
Ser Lys Leu Glu Val Thr Phe Asn Leu Arg Asp Val Thr Gly Gly
```

```
                710                 715                 720
Leu His Ser Ser Pro Arg Gln Ala Pro Gly Ser Gln Asp His Gln
                725                 730                 735
Gly Phe Phe Leu Ser Asp Asn Gln Ser Asp Ala Ala Pro Pro Leu
                740                 745                 750
Ser Ile Ser Asp Ala Phe Trp Leu Trp Pro Glu Leu Leu Gln Glu
                755                 760                 765
Met Pro Pro Lys His Ser Pro Gln Ser Pro Gln Glu Asp Pro Asp
                770                 775                 780
Cys Trp Pro Leu Lys Leu Gly Ser Arg Leu Glu Gln Leu Gln Ala
                785                 790                 795
Gln Met Asn Arg Leu Glu Ser Arg Val Ser Ser Asp Leu Ser Arg
                800                 805                 810
Ile Leu Gln Leu Leu Gln Lys Pro Met Pro Gln Gly His Ala Ser
                815                 820                 825
Tyr Ile Leu Glu Ala Pro Ala Ser Asn Asp Leu Ala Leu Val Pro
                830                 835                 840
Ile Ala Ser Glu Thr Thr Ser Pro Gly Pro Arg Leu Pro Gln Gly
                845                 850                 855
Phe Leu Pro Pro Ala Gln Thr Pro Ser Tyr Gly Asp Leu Asp Asp
                860                 865                 870
Cys Ser Pro Lys His Arg Asn Ser Ser Pro Arg Met Pro His Leu
                875                 880                 885
Ala Val Ala Met Asp Lys Thr Leu Ala Pro Ser Ser Glu Gln Glu
                890                 895                 900
Gln Pro Glu Gly Leu Trp Pro Pro Leu Ala Ser Pro Leu His Pro
                905                 910                 915
Leu Glu Val Gln Gly Leu Ile Cys Gly Pro Cys Phe Ser Ser Leu
                920                 925                 930
Pro Glu His Leu Gly Ser Val Pro Lys Gln Leu Asp Phe Gln Arg
                935                 940                 945
His Gly Ser Asp Pro Gly Phe Ala Gly Ser Trp Gly His
                950                 955
```

<210> 10
<211> 724
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472728CD1

<400> 10

```
Met Gly His Gln Gly Pro Phe Glu Glu Gly Asn Gly Gly Leu Arg
  1                 5                  10                  15
Val Ile Ala Thr Trp Arg Arg Lys Glu Ala Trp Arg Arg Asp Cys
                 20                  25                  30
Leu Leu Gly Ala Leu Pro Ser Val Ser Cys Gly Gly Trp Gly His
                 35                  40                  45
Arg Gly Arg Gln Thr Tyr Gly Arg Ala Cys Gly Val Lys Glu Lys
                 50                  55                  60
Pro Phe Ser Leu Leu Gly Pro Gln Ile Thr Val Tyr Ala Val Trp
                 65                  70                  75
Pro Gln Ser Glu Gly Pro Gln Glu Gly Arg Leu Arg Val Asn Ser
                 80                  85                  90
Ala Cys Leu Pro Pro Glu Arg Gly Leu Thr Asn Ala Cys Thr Asn
                 95                 100                 105
His Glu Glu Leu Ser Leu Asp Cys Leu Leu Phe Glu Asn Val Asn
                110                 115                 120
Thr Leu Thr Leu Asp Phe Cys Leu Trp Glu Lys Thr Thr Ile Val
                125                 130                 135
Pro Gly Val Leu Pro Tyr Ala Gly Leu Thr Leu Gln Ser Lys Phe
                140                 145                 150
Leu Leu Gly Arg Ala Leu Leu Ala Gly Val His Val Ile Thr Leu
                155                 160                 165
Thr Pro Glu Arg Val Thr His His Val His Gly Trp Tyr Met Glu
```

```
                170                 175                 180
Asp Gly Phe Lys Gly Asp Arg Thr Glu Gly Cys Arg Ser Asp Ser
            185                 190                 195
Val Ala Val Pro Ala Ala Ala Pro Val Cys Gln Pro Lys Ser Ala
        200                 205                 210
Thr Asn Gly Gln Pro Pro Ala Pro Ala Pro Thr Pro Thr Pro Arg
    215                 220                 225
Leu Ser Ile Ser Ser Arg Ala Thr Val Val Ala Arg Met Glu Gly
230                 235                 240
Thr Ser Gln Gly Gly Leu Gln Thr Val Met Lys Trp Lys Thr Val
                245                 250                 255
Val Ala Ile Phe Val Val Val Val Val Tyr Leu Val Thr Gly Gly
            260                 265                 270
Leu Val Phe Arg Ala Leu Glu Gln Pro Phe Glu Ser Ser Gln Lys
        275                 280                 285
Asn Thr Ile Ala Leu Glu Lys Ala Glu Phe Leu Arg Asp His Val
    290                 295                 300
Cys Val Ser Pro Gln Glu Leu Glu Thr Leu Ile Gln His Ala Leu
305                 310                 315
Asp Ala Asp Asn Ala Gly Val Ser Pro Ile Gly Asn Ser Ser Asn
                320                 325                 330
Asn Ser Ser His Trp Asp Leu Gly Ser Ala Phe Phe Phe Ala Gly
            335                 340                 345
Thr Val Ile Thr Thr Met Tyr Gly Asn Ile Ala Pro Ser Thr Glu
        350                 355                 360
Gly Gly Lys Ile Phe Cys Ile Leu Tyr Ala Ile Phe Gly Ile Pro
    365                 370                 375
Leu Phe Gly Phe Leu Leu Ala Gly Ile Gly Asp Gln Leu Gly Thr
380                 385                 390
Ile Phe Gly Lys Ser Ile Ala Arg Val Glu Lys Val Phe Arg Lys
                395                 400                 405
Lys Gln Val Ser Gln Thr Lys Ile Arg Val Ile Ser Thr Ile Leu
            410                 415                 420
Phe Ile Leu Ala Gly Cys Ile Val Phe Val Thr Ile Pro Ala Val
        425                 430                 435
Ile Phe Lys Tyr Ile Glu Gly Trp Thr Ala Leu Glu Ser Ile Tyr
    440                 445                 450
Phe Val Val Val Thr Leu Thr Thr Val Gly Phe Gly Asp Phe Val
455                 460                 465
Ala Val Val Val Phe Arg Gly Asn Ala Gly Ile Asn Tyr Arg Glu
                470                 475                 480
Trp Tyr Lys Pro Leu Val Trp Phe Trp Ile Leu Val Gly Leu Ala
            485                 490                 495
Tyr Phe Ala Ala Val Leu Ser Met Ile Gly Asp Trp Leu Arg Val
        500                 505                 510
Leu Ser Lys Lys Thr Lys Glu Glu Val Gly Glu Ile Lys Ala His
    515                 520                 525
Ala Ala Glu Trp Lys Ala Asn Val Thr Ala Glu Phe Arg Glu Thr
530                 535                 540
Arg Arg Arg Leu Ser Val Glu Ile His Asp Lys Leu Gln Arg Ala
                545                 550                 555
Ala Thr Ile Arg Ser Met Glu Arg Arg Arg Leu Gly Leu Asp Gln
            560                 565                 570
Arg Ala His Ser Leu Asp Met Leu Ser Pro Glu Lys Arg Ser Val
        575                 580                 585
Phe Ala Ala Leu Asp Thr Gly Arg Phe Lys Ala Ser Ser Gln Glu
    590                 595                 600
Ser Ile Asn Asn Arg Pro Asn Asn Leu Arg Leu Lys Gly Pro Glu
605                 610                 615
Gln Leu Asn Lys His Gly Gln Gly Ala Ser Glu Asp Asn Ile Ile
                620                 625                 630
Asn Lys Phe Gly Ser Thr Ser Arg Leu Thr Lys Arg Lys Asn Lys
            635                 640                 645
Asp Leu Lys Lys Thr Leu Pro Glu Asp Val Gln Lys Ile Tyr Lys
        650                 655                 660
Thr Phe Arg Asn Tyr Ser Leu Asp Glu Glu Lys Lys Glu Glu Glu
    665                 670                 675
```

```
Thr Glu Lys Met Cys Asn Ser Asp Asn Ser Ser Thr Ala Met Leu
                680                 685                 690
Thr Asp Cys Ile Gln Gln His Ala Glu Leu Glu Asn Gly Met Ile
                695                 700                 705
Pro Thr Asp Thr Lys Asp Arg Glu Pro Glu Asn Asn Ser Leu Leu
                710                 715                 720
Glu Asp Arg Asn
```

<210> 11
<211> 470
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474322CD1

<220>
<221> unsure
<222> 253
<223> unknown or other

<400> 11

```
Met Tyr Asn Glu Ile Leu Met Leu Gly Ala Lys Leu His Pro Thr
  1               5                  10                  15
Leu Lys Leu Glu Glu Leu Thr Asn Lys Lys Gly Met Thr Pro Leu
                20                  25                  30
Ala Leu Ala Ala Gly Thr Gly Lys Ile Gly Asn Arg His Asp Met
                35                  40                  45
Leu Leu Val Glu Pro Leu Asn Arg Leu Leu Gln Asp Lys Trp Asp
                50                  55                  60
Arg Phe Val Lys Arg Ile Phe Tyr Phe Asn Phe Leu Val Tyr Cys
                65                  70                  75
Leu Tyr Met Ile Ile Phe Thr Met Ala Ala Tyr Tyr Arg Pro Val
                80                  85                  90
Asp Gly Leu Pro Pro Phe Lys Met Glu Lys Thr Gly Asp Tyr Phe
                95                 100                 105
Arg Val Thr Gly Glu Ile Leu Ser Val Leu Gly Gly Val Tyr Phe
               110                 115                 120
Phe Phe Arg Gly Ile Gln Tyr Phe Leu Gln Arg Arg Pro Ser Met
               125                 130                 135
Lys Thr Leu Phe Val Asp Ser Tyr Ser Glu Met Leu Leu Phe Leu
               140                 145                 150
Gln Ser Leu Phe Met Leu Ala Thr Val Val Leu Tyr Phe Ser His
               155                 160                 165
Leu Lys Glu Tyr Val Ala Ser Met Val Phe Ser Leu Ala Leu Gly
               170                 175                 180
Trp Thr Asn Met Leu Tyr Tyr Thr Arg Gly Phe Gln Gln Met Gly
               185                 190                 195
Ile Tyr Ala Val Met Ile Glu Lys Met Ile Leu Arg Asp Leu Cys
               200                 205                 210
Arg Phe Met Phe Val Tyr Ile Val Phe Leu Phe Gly Phe Ser Thr
               215                 220                 225
Ala Val Val Thr Leu Ile Glu Asp Gly Lys Asn Asp Ser Leu Pro
               230                 235                 240
Ser Glu Ser Thr Ser His Arg Trp Arg Gly Pro Ala Xaa Arg Pro
               245                 250                 255
Asn Ser Ser Tyr Asn Ser Leu Tyr Ser Thr Cys Leu Glu Leu Phe
               260                 265                 270
Lys Phe Thr Ile Gly Met Gly Asp Leu Glu Phe Thr Glu Asn Tyr
               275                 280                 285
Asp Phe Lys Ala Val Phe Ile Ile Leu Leu Leu Ala Tyr Val Ile
               290                 295                 300
Leu Thr Tyr Ile Val Leu Leu Leu Asn Met Leu Ile Ala Leu Met
               305                 310                 315
Gly Glu Thr Val Glu Asn Val Ser Lys Glu Ser Glu Arg Ile Trp
```

```
                320                 325                 330
Arg Leu Gln Arg Ala Ile Thr Ile Leu Asp Thr Glu Lys Ser Phe
                335                 340                 345
Leu Lys Cys Met Arg Lys Ala Phe Arg Ser Gly Lys Leu Leu Gln
                350                 355                 360
Val Gly Tyr Thr Pro Asp Gly Lys Asp Asp Tyr Arg Trp Cys Phe
                365                 370                 375
Val Asp Glu Val Asn Trp Thr Thr Trp Asn Thr Asn Val Gly Ile
                380                 385                 390
Ile Asn Glu Asp Pro Gly Asn Cys Glu Gly Val Lys Arg Thr Leu
                395                 400                 405
Ser Phe Ser Leu Arg Ser Ser Arg Val Ser Gly Arg His Trp Lys
                410                 415                 420
Asn Phe Ala Leu Val Pro Leu Leu Arg Glu Ala Ser Ala Arg Asp
                425                 430                 435
Arg Gln Ser Ala Gln Pro Glu Glu Val Tyr Leu Arg Gln Phe Ser
                440                 445                 450
Gly Ser Leu Lys Pro Glu Asp Ala Glu Val Phe Lys Ser Pro Ala
                455                 460                 465
Ala Ser Gly Glu Lys
            470
```

<210> 12
<211> 618
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5455621CD1

<400> 12

```
Met Glu Val Lys Asn Phe Ala Val Trp Asp Tyr Val Val Phe Ala
  1                 5                  10                  15
Ala Leu Phe Phe Ile Ser Ser Gly Ile Gly Val Phe Phe Ala Ile
                 20                  25                  30
Lys Glu Arg Lys Lys Ala Thr Ser Arg Glu Phe Leu Val Gly Gly
                 35                  40                  45
Arg Gln Met Ser Phe Gly Pro Val Gly Leu Ser Leu Thr Ala Ser
                 50                  55                  60
Phe Met Ser Ala Val Thr Val Leu Gly Thr Pro Ser Glu Val Tyr
                 65                  70                  75
Arg Phe Gly Ala Ser Phe Leu Val Phe Phe Ile Ala Tyr Leu Phe
                 80                  85                  90
Val Ile Leu Leu Thr Ser Glu Leu Phe Leu Pro Val Phe Tyr Arg
                 95                 100                 105
Ser Gly Ile Thr Ser Thr Tyr Glu Tyr Leu Gln Leu Arg Phe Asn
                110                 115                 120
Lys Pro Val Arg Tyr Ala Ala Thr Val Ile Tyr Ile Val Gln Thr
                125                 130                 135
Ile Leu Tyr Thr Gly Val Val Val Tyr Ala Pro Ala Leu Ala Leu
                140                 145                 150
Asn Gln Val Thr Gly Phe Asp Leu Trp Gly Ser Val Phe Ala Thr
                155                 160                 165
Gly Ile Val Cys Thr Phe Tyr Cys Thr Leu Gly Gly Leu Lys Ala
                170                 175                 180
Val Val Trp Thr Asp Ala Phe Gln Met Val Val Met Ile Val Gly
                185                 190                 195
Phe Leu Thr Val Leu Ile Gln Gly Ser Thr His Ala Gly Gly Phe
                200                 205                 210
His Asn Val Leu Glu Gln Ser Thr Asn Gly Ser Arg Leu His Ile
                215                 220                 225
Phe Asp Phe Asp Val Asp Pro Leu Arg Arg His Thr Phe Trp Thr
                230                 235                 240
Ile Thr Val Gly Gly Thr Phe Thr Trp Leu Gly Ile Tyr Gly Val
                245                 250                 255
Asn Gln Ser Thr Ile Gln Arg Cys Ile Ser Cys Lys Thr Glu Lys
```

```
                260                 265                 270
His Ala Lys Leu Ala Leu Tyr Phe Asn Leu Leu Gly Leu Trp Ile
            275                 280                 285
Ile Leu Val Cys Ala Val Phe Ser Gly Leu Ile Met Tyr Ser His
        290                 295                 300
Phe Lys Asp Cys Asp Pro Trp Thr Ser Gly Ile Ile Ser Ala Pro
    305                 310                 315
Asp Gln Leu Met Pro Tyr Phe Val Met Glu Ile Phe Ala Thr Met
                320                 325                 330
Pro Gly Leu Pro Gly Leu Phe Val Ala Cys Ala Phe Ser Gly Thr
            335                 340                 345
Leu Ser Thr Val Ala Ser Ser Ile Asn Ala Leu Ala Thr Val Thr
        350                 355                 360
Phe Glu Asp Phe Val Lys Ser Cys Phe Pro His Leu Ser Asp Lys
    365                 370                 375
Leu Ser Thr Trp Ile Ser Lys Gly Leu Cys Leu Leu Phe Gly Val
                380                 385                 390
Met Cys Thr Ser Met Ala Val Ala Ala Ser Val Met Gly Gly Val
            395                 400                 405
Val Gln Ala Ser Leu Ser Ile His Gly Met Cys Gly Gly Pro Met
        410                 415                 420
Leu Gly Leu Phe Ser Leu Gly Ile Val Phe Pro Phe Val Asn Trp
    425                 430                 435
Lys Gly Ala Leu Gly Gly Leu Leu Thr Gly Ile Thr Leu Ser Phe
                440                 445                 450
Trp Val Ala Ile Gly Ala Phe Ile Tyr Pro Ala Pro Ala Ser Lys
            455                 460                 465
Thr Trp Pro Leu Pro Leu Ser Thr Asp Gln Cys Ile Lys Ser Asn
        470                 475                 480
Val Thr Ala Thr Gly Pro Pro Val Leu Ser Ser Arg Pro Gly Ile
    485                 490                 495
Ala Asp Thr Trp Tyr Ser Ile Ser Tyr Leu Tyr Tyr Ser Ala Leu
                500                 505                 510
Gly Cys Leu Gly Cys Ile Val Ala Gly Val Ile Ile Ser Leu Ile
            515                 520                 525
Thr Gly Arg Gln Arg Gly Glu Asp Ile Gln Pro Leu Leu Ile Arg
        530                 535                 540
Pro Val Cys Asn Leu Phe Cys Phe Trp Ser Lys Lys Tyr Lys Thr
    545                 550                 555
Leu Cys Trp Cys Gly Val Gln His Asp Ser Gly Thr Glu Gln Glu
                560                 565                 570
Asn Leu Glu Asn Gly Ser Ala Arg Lys Gln Gly Ala Glu Ser Val
            575                 580                 585
Leu Gln Asn Gly Leu Arg Arg Glu Ser Leu Val His Val Pro Gly
        590                 595                 600
Tyr Asp Pro Lys Asp Lys Ser Tyr Asn Asn Met Ala Phe Glu Thr
    605                 610                 615
Thr His Phe
```

<210> 13
<211> 631
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477248CD1

<400> 13

```
Met Glu Arg Gln Ser Arg Val Met Ser Glu Lys Asp Glu Tyr Gln
  1               5                  10                  15
Phe Gln His Gln Gly Ala Val Glu Leu Leu Val Phe Asn Phe Leu
            20                  25                  30
Leu Ile Leu Thr Ile Leu Thr Ile Trp Leu Phe Lys Asn His Arg
        35                  40                  45
Phe Arg Phe Leu His Glu Thr Gly Gly Ala Met Val Tyr Gly Leu
```

```
                50                  55                  60
Ile Met Gly Leu Ile Leu Arg Tyr Ala Thr Ala Pro Thr Asp Ile
            65                  70                  75
Glu Ser Gly Thr Val Tyr Asp Cys Val Lys Leu Thr Phe Ser Pro
            80                  85                  90
Ser Thr Leu Leu Val Asn Ile Thr Asp Gln Val Tyr Glu Tyr Lys
            95                 100                 105
Tyr Lys Arg Glu Ile Ser Gln His Asn Ile Asn Pro His Gln Gly
           110                 115                 120
Asn Ala Ile Leu Glu Lys Met Thr Phe Asp Pro Glu Ile Phe Phe
           125                 130                 135
Asn Val Leu Leu Pro Pro Ile Ile Phe His Ala Gly Tyr Ser Leu
           140                 145                 150
Lys Lys Arg His Phe Phe Gln Asn Leu Gly Ser Ile Leu Thr Tyr
           155                 160                 165
Ala Phe Leu Gly Thr Ala Ile Ser Cys Ile Val Ile Gly Leu Ile
           170                 175                 180
Met Tyr Gly Phe Val Lys Ala Met Ile His Ala Gly Gln Leu Lys
           185                 190                 195
Asn Gly Asp Phe His Phe Thr Asp Cys Leu Phe Phe Gly Ser Leu
           200                 205                 210
Met Ser Ala Thr Asp Pro Val Thr Val Leu Ala Ile Phe His Glu
           215                 220                 225
Leu His Val Asp Pro Asp Leu Tyr Thr Leu Leu Phe Gly Glu Ser
           230                 235                 240
Val Leu Asn Asp Ala Val Ala Ile Val Leu Thr Tyr Ser Ile Ser
           245                 250                 255
Ile Tyr Ser Pro Lys Glu Asn Pro Asn Ala Phe Asp Ala Ala Ala
           260                 265                 270
Phe Phe Gln Ser Val Gly Asn Phe Leu Gly Ile Phe Ala Gly Ser
           275                 280                 285
Phe Ala Met Gly Ser Ala Tyr Ala Ile Ile Thr Ala Leu Leu Thr
           290                 295                 300
Lys Phe Thr Lys Leu Cys Glu Phe Pro Met Leu Glu Thr Gly Leu
           305                 310                 315
Phe Phe Leu Leu Ser Trp Ser Ala Phe Leu Ser Ala Glu Ala Ala
           320                 325                 330
Gly Leu Thr Gly Ile Val Ala Val Leu Phe Cys Gly Val Thr Gln
           335                 340                 345
Ala His Tyr Thr Tyr Asn Asn Leu Ser Ser Asp Ser Lys Ile Arg
           350                 355                 360
Thr Lys Gln Leu Phe Glu Phe Met Asn Phe Leu Ala Glu Asn Val
           365                 370                 375
Ile Phe Cys Tyr Met Gly Leu Ala Leu Phe Thr Phe Gln Asn His
           380                 385                 390
Ile Phe Asn Ala Leu Phe Ile Leu Gly Ala Phe Leu Ala Ile Phe
           395                 400                 405
Val Ala Arg Ala Cys Asn Ile Tyr Pro Leu Ser Phe Leu Leu Asn
           410                 415                 420
Leu Gly Arg Lys Gln Lys Ile Pro Trp Asn Phe Gln His Met Met
           425                 430                 435
Met Phe Ser Gly Leu Arg Gly Ala Ile Ala Phe Ala Leu Ala Ile
           440                 445                 450
Arg Asn Thr Glu Ser Gln Pro Lys Gln Met Met Phe Thr Thr Thr
           455                 460                 465
Leu Leu Leu Val Phe Phe Thr Val Trp Val Phe Gly Gly Gly Thr
           470                 475                 480
Thr Pro Met Leu Thr Trp Leu Gln Ile Arg Val Gly Val Asp Leu
           485                 490                 495
Asp Glu Asn Leu Lys Glu Asp Pro Ser Ser Gln His Gln Glu Ala
           500                 505                 510
Asn Asn Leu Asp Lys Asn Met Thr Lys Ala Glu Ser Ala Arg Leu
           515                 520                 525
Phe Arg Met Trp Tyr Ser Phe Asp His Lys Tyr Leu Lys Pro Ile
           530                 535                 540
Leu Thr His Ser Gly Pro Pro Leu Thr Thr Thr Leu Pro Glu Trp
           545                 550                 555
```

```
Cys Gly Pro Ile Ser Arg Leu Leu Thr Ser Pro Gln Ala Tyr Gly
                560                 565                 570
Glu Gln Leu Lys Glu Asp Asp Val Glu Cys Ile Val Asn Gln Asp
                575                 580                 585
Glu Leu Ala Ile Asn Tyr Gln Glu Gln Ala Ser Ser Pro Cys Ser
                590                 595                 600
Pro Pro Ala Arg Leu Gly Leu Asp Gln Lys Ala Ser Pro Gln Thr
                605                 610                 615
Pro Gly Lys Glu Asn Ile Tyr Glu Gly Asp Leu Gly Pro Gly Arg
                620                 625                 630
Leu
```

<210> 14
<211> 1256
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2944004CD1

<400> 14

```
Met Asp Arg Glu Glu Arg Lys Thr Ile Asn Gln Gly Gln Glu Asp
  1                 5                  10                  15
Glu Met Glu Ile Tyr Gly Tyr Asn Leu Ser Arg Trp Lys Leu Ala
                 20                  25                  30
Ile Val Ser Leu Gly Val Ile Cys Ser Gly Gly Val Ser Pro Pro
                 35                  40                  45
Pro Leu Tyr Trp Met Pro Glu Trp Arg Val Lys Ala Thr Cys Val
                 50                  55                  60
Arg Ala Ala Ile Lys Asp Cys Glu Val Val Leu Leu Arg Thr Thr
                 65                  70                  75
Asp Glu Phe Lys Met Trp Phe Cys Ala Lys Ile Arg Val Leu Ser
                 80                  85                  90
Leu Glu Thr Tyr Pro Val Ser Ser Pro Lys Ser Met Ser Asn Lys
                 95                 100                 105
Leu Ser Asn Gly His Ala Val Cys Leu Ile Glu Asn Pro Thr Glu
                110                 115                 120
Glu Asn Arg His Arg Ile Ser Lys Tyr Ser Gln Thr Glu Ser Gln
                125                 130                 135
Gln Ile Arg Tyr Phe Thr His His Ser Val Lys Tyr Phe Trp Asn
                140                 145                 150
Asp Thr Ile His Asn Phe Asp Phe Leu Lys Gly Leu Asp Glu Gly
                155                 160                 165
Val Ser Cys Thr Ser Ile Tyr Glu Lys His Ser Ala Gly Leu Thr
                170                 175                 180
Lys Gly Met His Ala Tyr Arg Lys Leu Leu Tyr Gly Val Asn Glu
                185                 190                 195
Ile Ala Val Lys Val Pro Ser Val Phe Lys Leu Leu Ile Lys Glu
                200                 205                 210
Val Leu Asn Pro Phe Tyr Ile Phe Gln Leu Phe Ser Val Ile Leu
                215                 220                 225
Trp Ser Thr Asp Glu Tyr Tyr Tyr Tyr Ala Leu Ala Ile Val Val
                230                 235                 240
Met Ser Ile Val Ser Ile Val Ser Ser Leu Tyr Ser Ile Arg Lys
                245                 250                 255
Gln Tyr Val Met Leu His Asp Met Val Ala Thr His Ser Thr Val
                260                 265                 270
Arg Val Ser Val Cys Arg Val Asn Glu Glu Ile Glu Glu Ile Phe
                275                 280                 285
Ser Thr Asp Leu Val Pro Gly Asp Val Met Val Ile Pro Leu Asn
                290                 295                 300
Gly Thr Ile Met Pro Cys Asp Ala Val Leu Ile Asn Gly Thr Cys
                305                 310                 315
Ile Val Asn Glu Ser Met Leu Thr Gly Glu Ser Val Pro Val Thr
                320                 325                 330
```

```
Lys Thr Asn Leu Pro Asn Pro Ser Val Asp Val Lys Gly Ile Gly
                335                 340                 345
Asp Glu Leu Tyr Asn Pro Glu Thr His Lys Arg His Thr Leu Phe
            350                 355                 360
Cys Gly Thr Thr Val Ile Gln Thr Arg Phe Tyr Thr Gly Glu Leu
        365                 370                 375
Val Lys Ala Ile Val Val Arg Thr Gly Phe Ser Thr Ser Lys Gly
    380                 385                 390
Gln Leu Val Arg Ser Ile Leu Tyr Pro Lys Pro Thr Asp Phe Lys
395                 400                 405
Leu Tyr Arg Asp Ala Tyr Leu Phe Leu Leu Cys Leu Val Ala Val
            410                 415                 420
Ala Gly Ile Gly Phe Ile Tyr Thr Ile Ile Asn Ser Ile Leu Asn
        425                 430                 435
Glu Val Gln Val Gly Val Ile Ile Ile Glu Ser Leu Asp Ile Ile
    440                 445                 450
Thr Ile Thr Val Pro Pro Ala Leu Pro Ala Ala Met Thr Ala Gly
455                 460                 465
Ile Val Tyr Ala Gln Arg Arg Leu Lys Lys Ile Gly Ile Phe Cys
            470                 475                 480
Ile Ser Pro Gln Arg Ile Asn Ile Cys Gly Gln Leu Asn Leu Val
        485                 490                 495
Cys Phe Asp Lys Thr Gly Thr Leu Thr Glu Asp Gly Leu Asp Leu
    500                 505                 510
Trp Gly Ile Gln Arg Val Glu Asn Ala Arg Phe Leu Ser Pro Glu
515                 520                 525
Glu Asn Val Cys Asn Glu Met Leu Val Lys Ser Gln Phe Val Ala
            530                 535                 540
Cys Met Ala Thr Cys His Ser Leu Thr Lys Ile Glu Gly Val Leu
        545                 550                 555
Ser Gly Asp Pro Leu Asp Leu Lys Met Phe Glu Ala Ile Gly Trp
    560                 565                 570
Ile Leu Glu Glu Ala Thr Glu Glu Glu Thr Ala Leu His Asn Arg
575                 580                 585
Ile Met Pro Thr Val Val Arg Pro Pro Lys Gln Leu Leu Pro Glu
            590                 595                 600
Ser Thr Pro Ala Gly Asn Gln Glu Met Glu Leu Phe Glu Leu Pro
        605                 610                 615
Ala Thr Tyr Glu Ile Gly Ile Val Arg Gln Phe Pro Phe Ser Ser
    620                 625                 630
Ala Leu Gln Arg Met Ser Val Val Ala Arg Val Leu Gly Asp Arg
635                 640                 645
Lys Met Asp Ala Tyr Met Lys Gly Ala Pro Glu Ala Ile Ala Gly
            650                 655                 660
Leu Cys Lys Pro Glu Thr Val Pro Val Asp Phe Gln Asn Val Leu
        665                 670                 675
Glu Asp Phe Thr Lys Gln Gly Phe Arg Val Ile Ala Leu Ala His
    680                 685                 690
Arg Lys Leu Glu Ser Lys Leu Thr Trp His Lys Val Gln Asn Ile
695                 700                 705
Ser Arg Asp Ala Ile Glu Asn Asn Met Asp Phe Met Gly Leu Ile
            710                 715                 720
Ile Met Gln Asn Lys Leu Lys Gln Lys Thr Pro Ala Val Leu Glu
        725                 730                 735
Asp Leu His Lys Ala Asn Ile Arg Thr Val Met Val Thr Gly Asp
    740                 745                 750
Ser Met Leu Thr Ala Val Ser Val Ala Arg Asp Cys Gly Met Ile
755                 760                 765
Leu Pro Gln Asp Lys Val Ile Ile Ala Glu Ala Leu Pro Pro Lys
            770                 775                 780
Asp Gly Lys Val Ala Lys Ile Asn Trp His Tyr Ala Asp Ser Leu
        785                 790                 795
Thr Gln Cys Ser His Pro Ser Ala Ile Asp Pro Glu Ala Ile Pro
    800                 805                 810
Val Lys Leu Val His Asp Ser Leu Glu Asp Leu Gln Met Thr Arg
815                 820                 825
Tyr His Phe Ala Met Asn Gly Lys Ser Phe Ser Val Ile Leu Glu
```

```
                830                 835                 840
His Phe Gln Asp Leu Val Pro Lys Leu Met Leu His Gly Thr Val
                845                 850                 855
Phe Ala Arg Met Ala Pro Asp Gln Lys Thr Gln Leu Ile Glu Ala
                860                 865                 870
Leu Gln Asn Val Asp Tyr Phe Val Gly Met Cys Gly Asp Gly Ala
                875                 880                 885
Asn Asp Cys Gly Ala Leu Lys Arg Ala His Gly Gly Ile Ser Leu
                890                 895                 900
Ser Glu Leu Glu Ala Ser Val Ala Ser Pro Phe Thr Ser Lys Thr
                905                 910                 915
Pro Ser Ile Ser Cys Val Pro Asn Leu Ile Arg Glu Gly Arg Ala
                920                 925                 930
Ala Leu Ile Thr Ser Phe Cys Val Phe Lys Phe Met Ala Leu Tyr
                935                 940                 945
Ser Ile Ile Gln Tyr Phe Ser Val Thr Leu Leu Tyr Ser Ile Leu
                950                 955                 960
Ser Asn Leu Gly Asp Phe Gln Phe Leu Phe Ile Asp Leu Ala Ile
                965                 970                 975
Ile Leu Val Val Val Phe Thr Met Ser Leu Asn Pro Ala Trp Lys
                980                 985                 990
Glu Leu Val Ala Gln Arg Pro Pro Ser Gly Leu Ile Ser Gly Ala
                995                1000                1005
Leu Leu Phe Ser Val Leu Ser Gln Ile Ile Ile Cys Ile Gly Phe
               1010                1015                1020
Gln Ser Leu Gly Phe Phe Trp Val Lys Gln Gln Pro Trp Tyr Glu
               1025                1030                1035
Val Trp His Pro Lys Ser Asp Ala Cys Asn Thr Thr Gly Ser Gly
               1040                1045                1050
Phe Trp Asn Ser Ser His Val Asp Asn Glu Thr Glu Leu Asp Glu
               1055                1060                1065
His Asn Ile Gln Asn Tyr Glu Asn Thr Thr Val Phe Phe Ile Ser
               1070                1075                1080
Ser Phe Gln Tyr Leu Ile Val Ala Ile Ala Phe Ser Lys Gly Lys
               1085                1090                1095
Pro Phe Arg Gln Pro Cys Tyr Lys Asn Tyr Phe Phe Val Phe Ser
               1100                1105                1110
Val Ile Phe Leu Tyr Ile Phe Ile Leu Phe Ile Met Leu Tyr Pro
               1115                1120                1125
Val Ala Ser Val Asp Gln Val Leu Gln Ile Val Cys Val Pro Tyr
               1130                1135                1140
Gln Trp Arg Val Thr Met Leu Ile Ile Val Leu Val Asn Ala Phe
               1145                1150                1155
Val Ser Ile Thr Val Glu Asn Phe Phe Leu Asp Met Val Leu Trp
               1160                1165                1170
Lys Val Val Phe Asn Arg Asp Lys Gln Gly Glu Tyr Arg Phe Ser
               1175                1180                1185
Thr Thr Gln Pro Pro Gln Glu Ser Val Asp Arg Trp Gly Lys Cys
               1190                1195                1200
Cys Leu Pro Trp Ala Leu Gly Cys Arg Lys Lys Thr Pro Lys Ala
               1205                1210                1215
Lys Tyr Met Tyr Leu Ala Gln Glu Leu Leu Val Asp Pro Glu Trp
               1220                1225                1230
Pro Pro Lys Pro Gln Thr Thr Thr Glu Ala Lys Ala Leu Val Lys
               1235                1240                1245
Glu Asn Gly Ser Cys Gln Ile Ile Thr Ile Thr
               1250                1255
```

<210> 15
<211> 499
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3046849CD1

<400> 15

```
Met Leu His Ala Leu Leu Arg Ser Arg Thr Ile Gln Gly Arg Ile
  1               5                  10                  15
Leu Leu Leu Thr Ile Cys Ala Ala Gly Ile Gly Gly Thr Phe Gln
             20                  25                  30
Phe Gly Tyr Asn Leu Ser Ile Ile Asn Ala Pro Thr Leu His Ile
         35                  40                  45
Gln Glu Phe Thr Asn Glu Thr Trp Gln Ala Arg Thr Gly Glu Pro
     50                  55                  60
Leu Pro Asp His Leu Val Leu Leu Met Trp Ser Leu Ile Val Ser
 65                  70                  75
Leu Tyr Pro Leu Gly Gly Leu Phe Gly Ala Leu Leu Ala Gly Pro
             80                  85                  90
Leu Ala Ile Thr Leu Gly Arg Lys Lys Ser Leu Leu Val Asn Asn
         95                 100                 105
Ile Phe Val Val Ser Ala Ala Ile Leu Phe Gly Phe Ser Arg Lys
    110                 115                 120
Ala Gly Ser Phe Glu Met Ile Met Leu Gly Arg Leu Leu Val Gly
125                 130                 135
Val Asn Ala Gly Val Ser Met Asn Ile Gln Pro Met Tyr Leu Gly
            140                 145                 150
Glu Ser Ala Pro Lys Glu Leu Arg Gly Ala Val Ala Met Ser Ser
        155                 160                 165
Ala Ile Phe Thr Ala Leu Gly Ile Val Met Gly Gln Val Val Gly
    170                 175                 180
Leu Arg Glu Leu Leu Gly Gly Pro Gln Ala Trp Pro Leu Leu Leu
185                 190                 195
Ala Ser Cys Leu Val Pro Gly Ala Leu Gln Leu Ala Ser Leu Pro
            200                 205                 210
Leu Leu Pro Glu Ser Pro Arg Tyr Leu Leu Ile Asp Cys Gly Asp
        215                 220                 225
Thr Glu Ala Cys Leu Ala Ala Leu Arg Gln Leu Arg Gly Ser Gly
    230                 235                 240
Asp Leu Ala Gly Glu Leu Glu Glu Leu Glu Glu Glu Arg Ala Ala
245                 250                 255
Cys Gln Gly Cys Arg Ala Arg Arg Pro Trp Glu Leu Phe Gln His
            260                 265                 270
Arg Ala Leu Arg Arg Gln Val Thr Ser Leu Val Val Leu Gly Ser
        275                 280                 285
Ala Met Glu Leu Cys Gly Asn Asp Ser Val Tyr Ala Tyr Ala Ser
    290                 295                 300
Ser Val Phe Arg Lys Ala Gly Val Pro Glu Ala Lys Ile Gln Tyr
305                 310                 315
Ala Ile Ile Gly Thr Gly Ser Cys Glu Leu Leu Thr Ala Val Val
            320                 325                 330
Ser Cys Val Val Ile Glu Arg Val Gly Arg Arg Val Leu Leu Ile
        335                 340                 345
Gly Gly Tyr Ser Leu Met Thr Cys Trp Gly Ser Ile Phe Thr Val
    350                 355                 360
Ala Leu Cys Leu Gln Ser Ser Phe Pro Trp Thr Leu Tyr Leu Ala
365                 370                 375
Met Ala Cys Ile Phe Ala Phe Ile Leu Ser Phe Gly Ile Gly Pro
            380                 385                 390
Ala Gly Val Thr Gly Ile Leu Ala Thr Glu Leu Phe Asp Gln Met
        395                 400                 405
Ala Arg Pro Ala Ala Cys Met Val Cys Gly Ala Leu Met Trp Ile
    410                 415                 420
Met Leu Ile Leu Val Gly Leu Gly Phe Pro Phe Ile Met Glu Ala
425                 430                 435
Leu Ser His Phe Leu Tyr Val Pro Phe Leu Gly Val Cys Val Cys
            440                 445                 450
Gly Ala Ile Tyr Thr Gly Leu Phe Leu Pro Glu Thr Lys Gly Lys
        455                 460                 465
Thr Phe Gln Glu Ile Ser Lys Glu Leu His Arg Leu Asn Phe Pro
    470                 475                 480
Arg Arg Ala Gln Gly Pro Thr Trp Arg Ser Leu Glu Val Ile Gln
485                 490                 495
```

Ser Thr Glu Leu

<210> 16
<211> 596
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4538363CD1

<400> 16

```
Met Ala Ala Asn Ser Thr Ser Asp Leu His Thr Pro Gly Thr Gln
  1               5                  10                  15
Leu Ser Val Ala Asp Ile Ile Val Ile Thr Val Tyr Phe Ala Leu
             20                  25                  30
Asn Val Ala Val Gly Ile Trp Ser Ser Cys Arg Ala Ser Arg Asn
             35                  40                  45
Thr Val Asn Gly Tyr Phe Leu Ala Gly Arg Asp Met Thr Trp Trp
             50                  55                  60
Pro Ile Gly Ala Ser Leu Phe Ala Ser Ser Glu Gly Ser Gly Leu
             65                  70                  75
Phe Ile Gly Leu Ala Gly Ser Gly Ala Ala Gly Gly Leu Ala Val
             80                  85                  90
Ala Gly Phe Glu Trp Asn Ala Thr Tyr Val Leu Leu Ala Leu Ala
             95                 100                 105
Trp Val Phe Val Pro Ile Tyr Ile Ser Ser Glu Ile Val Thr Leu
            110                 115                 120
Pro Glu Tyr Ile Gln Lys Arg Tyr Gly Gly Gln Arg Ile Arg Met
            125                 130                 135
Tyr Leu Ser Val Leu Ser Leu Leu Leu Ser Val Phe Thr Lys Ile
            140                 145                 150
Ser Leu Asp Leu Tyr Ala Gly Ala Leu Phe Val His Ile Cys Leu
            155                 160                 165
Gly Trp Asn Phe Tyr Leu Ser Thr Ile Leu Thr Leu Gly Ile Thr
            170                 175                 180
Ala Leu Tyr Thr Ile Ala Gly Gly Leu Ala Ala Val Ile Tyr Thr
            185                 190                 195
Asp Ala Leu Gln Thr Leu Ile Met Val Val Gly Ala Val Ile Leu
            200                 205                 210
Thr Ile Lys Ala Phe Asp Gln Ile Gly Gly Tyr Gly Gln Leu Glu
            215                 220                 225
Ala Ala Tyr Ala Gln Ala Ile Pro Ser Arg Thr Ile Ala Asn Thr
            230                 235                 240
Thr Cys His Leu Pro Arg Thr Asp Ala Met His Met Phe Arg Asp
            245                 250                 255
Pro His Thr Gly Asp Leu Pro Trp Thr Gly Met Thr Phe Gly Leu
            260                 265                 270
Thr Ile Met Ala Thr Trp Tyr Trp Cys Thr Asp Gln Val Ile Val
            275                 280                 285
Gln Arg Ser Leu Ser Ala Arg Asp Leu Asn His Ala Lys Ala Gly
            290                 295                 300
Ser Ile Leu Ala Ser Tyr Leu Lys Met Leu Pro Met Gly Leu Ile
            305                 310                 315
Ile Met Pro Gly Met Ile Ser Arg Ala Leu Phe Pro Asp Asp Val
            320                 325                 330
Gly Cys Val Val Pro Ser Glu Cys Leu Arg Ala Cys Gly Ala Glu
            335                 340                 345
Val Gly Cys Ser Asn Ile Ala Tyr Pro Lys Leu Val Met Glu Leu
            350                 355                 360
Met Pro Ile Gly Leu Arg Gly Leu Met Ile Ala Val Met Leu Ala
            365                 370                 375
Ala Leu Met Ser Ser Leu Thr Ser Ile Phe Asn Ser Ser Ser Thr
            380                 385                 390
Leu Phe Thr Met Asp Ile Trp Arg Arg Leu Arg Pro Arg Ser Gly
            395                 400                 405
```

```
Glu Arg Glu Leu Leu Leu Val Gly Arg Leu Val Ile Val Ala Leu
                410                 415                 420
Ile Gly Val Ser Val Ala Trp Ile Pro Val Leu Gln Asp Ser Asn
                425                 430                 435
Ser Gly Gln Leu Phe Ile Tyr Met Gln Ser Val Thr Ser Ser Leu
                440                 445                 450
Ala Pro Pro Val Thr Ala Val Phe Val Leu Gly Val Phe Trp Arg
                455                 460                 465
Arg Ala Asn Glu Gln Gly Ala Phe Trp Gly Leu Ile Ala Gly Leu
                470                 475                 480
Val Val Gly Ala Thr Arg Leu Val Leu Glu Phe Leu Asn Pro Ala
                485                 490                 495
Pro Pro Cys Gly Glu Pro Asp Thr Arg Pro Ala Val Leu Gly Ser
                500                 505                 510
Ile His Tyr Leu His Phe Ala Val Ala Leu Phe Ala Leu Ser Gly
                515                 520                 525
Ala Val Val Val Ala Gly Ser Leu Leu Thr Pro Pro Pro Gln Ser
                530                 535                 540
Val Gln Ile Glu Asn Leu Thr Trp Trp Thr Leu Ala Gln Asp Val
                545                 550                 555
Pro Leu Gly Thr Lys Ala Gly Asp Gly Gln Thr Pro Gln Lys His
                560                 565                 570
Ala Phe Trp Ala Arg Val Cys Gly Phe Asn Ala Ile Leu Leu Met
                575                 580                 585
Cys Val Asn Ile Phe Phe Tyr Ala Tyr Phe Ala
                590                 595

<210> 17
<211> 1192
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6427460CD1

<400> 17
Met Asp Cys Ser Leu Val Arg Thr Leu Val His Arg Tyr Cys Ala
  1               5                  10                  15
Gly Glu Glu Asn Trp Val Asp Ser Arg Thr Ile Tyr Val Gly His
                 20                  25                  30
Arg Glu Pro Pro Pro Gly Ala Glu Ala Tyr Ile Pro Gln Arg Tyr
                 35                  40                  45
Pro Asp Asn Arg Ile Val Ser Ser Lys Tyr Thr Phe Trp Asn Phe
                 50                  55                  60
Ile Pro Lys Asn Leu Phe Glu Gln Phe Arg Arg Val Ala Asn Phe
                 65                  70                  75
Tyr Phe Leu Ile Ile Phe Leu Val Gln Leu Ile Ile Asp Thr Pro
                 80                  85                  90
Thr Ser Pro Val Thr Ser Gly Leu Pro Leu Phe Phe Val Ile Thr
                 95                 100                 105
Val Thr Ala Ile Lys Gln Gly Tyr Glu Asp Trp Leu Arg His Lys
                110                 115                 120
Ala Asp Asn Ala Met Asn Gln Cys Pro Val His Phe Ile Gln His
                125                 130                 135
Gly Lys Leu Val Arg Lys Gln Ser Arg Lys Leu Arg Val Gly Asp
                140                 145                 150
Ile Val Met Val Lys Glu Asp Glu Thr Phe Pro Cys Asp Leu Ile
                155                 160                 165
Phe Leu Ser Ser Asn Arg Gly Asp Gly Thr Cys His Val Thr Thr
                170                 175                 180
Ala Ser Leu Asp Gly Glu Ser Ser His Lys Thr His Tyr Ala Val
                185                 190                 195
Gln Asp Thr Lys Gly Phe His Thr Glu Glu Asp Ile Gly Gly Leu
                200                 205                 210
His Ala Thr Ile Glu Cys Glu Gln Pro Gln Pro Asp Leu Tyr Lys
                215                 220                 225
```

Phe Val Gly Arg Ile Asn Val Tyr Ser Asp Leu Asn Asp Pro Val
230 235 240
Val Arg Pro Leu Gly Ser Glu Asn Leu Leu Leu Arg Gly Ala Thr
245 250 255
Leu Lys Asn Thr Glu Lys Ile Phe Gly Val Ala Ile Tyr Thr Gly
260 265 270
Met Glu Thr Lys Met Ala Leu Asn Tyr Gln Ser Lys Ser Gln Lys
275 280 285
Arg Ser Ala Val Glu Lys Ser Met Asn Ala Phe Leu Ile Val Tyr
290 295 300
Leu Cys Ile Leu Ile Ser Lys Ala Leu Ile Asn Thr Val Leu Lys
305 310 315
Tyr Val Trp Gln Ser Glu Pro Phe Arg Asp Glu Pro Trp Tyr Asn
320 325 330
Gln Lys Thr Glu Ser Glu Arg Gln Arg Asn Leu Phe Leu Lys Ala
335 340 345
Phe Thr Asp Phe Leu Ala Phe Met Val Leu Phe Asn Tyr Ile Ile
350 355 360
Pro Val Ser Met Tyr Val Thr Val Glu Met Gln Lys Phe Leu Gly
365 370 375
Ser Tyr Phe Ile Thr Trp Asp Glu Asp Met Phe Asp Glu Glu Thr
380 385 390
Gly Glu Gly Pro Leu Val Asn Thr Ser Asp Leu Asn Glu Glu Leu
395 400 405
Gly Gln Val Glu Tyr Ile Phe Thr Asp Lys Thr Gly Thr Leu Thr
410 415 420
Glu Asn Asn Met Glu Phe Lys Glu Cys Cys Ile Glu Gly His Val
425 430 435
Tyr Val Pro His Val Ile Cys Asn Gly Gln Val Leu Pro Glu Ser
440 445 450
Ser Gly Ile Asp Met Ile Asp Ser Ser Pro Ser Val Asn Gly Arg
455 460 465
Glu Arg Glu Glu Leu Phe Phe Arg Ala Leu Cys Leu Cys His Thr
470 475 480
Val Gln Val Lys Asp Asp Asp Ser Val Asp Gly Pro Arg Lys Ser
485 490 495
Pro Asp Gly Gly Lys Ser Cys Val Tyr Ile Ser Ser Ser Pro Asp
500 505 510
Glu Val Ala Leu Val Glu Gly Val Gln Arg Leu Gly Phe Thr Tyr
515 520 525
Leu Arg Leu Lys Asp Asn Tyr Met Glu Ile Leu Asn Arg Glu Asn
530 535 540
His Ile Glu Arg Phe Glu Leu Leu Glu Ile Leu Ser Phe Asp Ser
545 550 555
Val Arg Arg Arg Met Ser Val Ile Val Lys Ser Ala Thr Gly Glu
560 565 570
Ile Tyr Leu Phe Cys Lys Gly Ala Asp Ser Ser Ile Phe Pro Arg
575 580 585
Val Ile Glu Gly Lys Val Asp Gln Ile Arg Ala Arg Val Glu Arg
590 595 600
Asn Ala Val Glu Gly Leu Arg Thr Leu Cys Val Ala Tyr Lys Arg
605 610 615
Leu Ile Gln Glu Glu Tyr Glu Gly Ile Cys Lys Leu Leu Gln Ala
620 625 630
Ala Lys Val Ala Leu Gln Asp Arg Glu Lys Lys Leu Ala Glu Ala
635 640 645
Tyr Glu Gln Ile Glu Lys Asp Leu Thr Leu Leu Gly Ala Thr Ala
650 655 660
Val Glu Asp Arg Leu Gln Glu Lys Ala Ala Asp Thr Ile Glu Ala
665 670 675
Leu Gln Lys Ala Gly Ile Lys Val Trp Val Leu Thr Gly Asp Lys
680 685 690
Met Glu Thr Ala Ala Ala Thr Cys Tyr Ala Cys Lys Leu Phe Arg
695 700 705
Arg Asn Thr Gln Leu Leu Glu Leu Thr Thr Lys Arg Ile Glu Glu
710 715 720
Gln Ser Leu His Asp Val Leu Phe Glu Leu Ser Lys Thr Val Leu

```
                725                 730                 735
Arg His Ser Gly Ser Leu Thr Arg Asp Asn Leu Ser Gly Leu Ser
            740                 745                 750
Ala Asp Met Gln Asp Tyr Gly Leu Ile Ile Asp Gly Ala Ala Leu
        755                 760                 765
Ser Leu Ile Met Lys Pro Arg Glu Asp Gly Ser Ser Gly Asn Tyr
    770                 775                 780
Arg Glu Leu Phe Leu Glu Ile Cys Arg Ser Cys Ser Ala Val Leu
785                 790                 795
Cys Cys Arg Met Ala Pro Leu Gln Lys Ala Gln Ile Val Lys Leu
            800                 805                 810
Ile Lys Phe Ser Lys Glu His Pro Ile Thr Leu Ala Ile Gly Asp
        815                 820                 825
Gly Ala Asn Asp Val Ser Met Ile Leu Glu Ala His Val Gly Ile
    830                 835                 840
Gly Val Ile Gly Lys Glu Gly Arg Gln Ala Ala Arg Asn Ser Asp
845                 850                 855
Tyr Ala Ile Pro Lys Phe Lys His Leu Lys Lys Met Leu Leu Val
            860                 865                 870
His Gly His Phe Tyr Tyr Ile Arg Ile Ser Glu Leu Val Gln Tyr
        875                 880                 885
Phe Phe Tyr Lys Asn Val Cys Phe Ile Phe Pro Gln Phe Leu Tyr
    890                 895                 900
Gln Phe Phe Cys Gly Phe Ser Gln Gln Thr Leu Tyr Asp Thr Ala
905                 910                 915
Tyr Leu Thr Leu Tyr Asn Ile Ser Phe Thr Ser Leu Pro Ile Leu
            920                 925                 930
Leu Tyr Ser Leu Met Glu Gln His Val Gly Ile Asp Val Leu Lys
        935                 940                 945
Arg Asp Pro Thr Leu Tyr Arg Asp Val Ala Lys Asn Ala Leu Leu
    950                 955                 960
Arg Trp Arg Val Phe Ile Tyr Trp Thr Leu Leu Gly Leu Phe Asp
965                 970                 975
Ala Leu Val Phe Phe Phe Gly Ala Tyr Phe Val Phe Glu Asn Thr
            980                 985                 990
Thr Val Thr Ser Asn Gly Gln Ile Phe Gly Asn Trp Thr Phe Gly
        995                1000                1005
Thr Leu Val Phe Thr Val Met Val Phe Thr Val Thr Leu Lys Leu
    1010                1015                1020
Ala Leu Asp Thr His Tyr Trp Thr Trp Ile Asn His Phe Val Ile
1025                1030                1035
Trp Gly Ser Leu Leu Phe Tyr Val Val Phe Ser Leu Leu Trp Gly
            1040                1045                1050
Gly Val Ile Trp Pro Phe Leu Asn Tyr Gln Arg Met Tyr Tyr Val
        1055                1060                1065
Phe Ile Gln Met Leu Ser Ser Gly Pro Ala Trp Leu Ala Ile Val
    1070                1075                1080
Leu Leu Val Thr Ile Ser Leu Leu Pro Asp Val Leu Lys Lys Val
1085                1090                1095
Leu Cys Arg Gln Leu Trp Pro Thr Ala Thr Glu Arg Val Gln Gln
            1100                1105                1110
Asn Gly Cys Ala Gln Pro Arg Asp Arg Asp Ser Glu Phe Thr Pro
        1115                1120                1125
Leu Ala Ser Leu Gln Ser Pro Gly Tyr Gln Ser Thr Cys Pro Ser
    1130                1135                1140
Ala Ala Trp Tyr Ser Ser His Ser Gln Gln Val Thr Leu Ala Ala
1145                1150                1155
Trp Lys Glu Lys Val Ser Thr Glu Pro Pro Pro Ile Leu Gly Gly
            1160                1165                1170
Ser His His His Cys Ser Ser Ile Pro Ser His Ser Cys Pro Arg
        1175                1180                1185
Ser Arg Val Gly Met Leu Val
    1190
```

<210> 18
<211> 638
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474127CD1

<400> 18

```
Met Gly Lys Ile Glu Asn Asn Glu Arg Val Ile Leu Asn Val Gly
  1               5                  10                  15
Gly Thr Arg His Glu Thr Tyr Arg Ser Thr Leu Lys Thr Leu Pro
             20                  25                  30
Gly Thr Arg Leu Ala Leu Leu Ala Ser Ser Glu Pro Pro Gly Asp
         35                  40                  45
Cys Leu Thr Thr Ala Gly Asp Lys Leu Gln Pro Ser Pro Pro Pro
     50                  55                  60
Leu Ser Pro Pro Pro Arg Ala Pro Pro Leu Ser Pro Gly Pro Gly
 65                  70                  75
Gly Cys Phe Glu Gly Gly Ala Gly Asn Cys Ser Ser Arg Gly Gly
             80                  85                  90
Arg Ala Ser Asp His Pro Gly Gly Gly Arg Glu Phe Phe Phe Asp
         95                 100                 105
Arg His Pro Gly Val Phe Ala Tyr Val Leu Asn Tyr Tyr Arg Thr
    110                 115                 120
Gly Lys Leu His Cys Pro Ala Asp Val Cys Gly Pro Leu Phe Glu
125                 130                 135
Glu Glu Leu Ala Phe Trp Gly Ile Asp Glu Thr Asp Val Glu Pro
            140                 145                 150
Cys Cys Trp Met Thr Tyr Arg Gln His Arg Asp Ala Glu Glu Ala
        155                 160                 165
Leu Asp Ile Phe Glu Thr Pro Asp Leu Ile Gly Gly Asp Pro Gly
    170                 175                 180
Asp Asp Glu Asp Leu Ala Ala Lys Arg Leu Gly Ile Glu Asp Ala
185                 190                 195
Ala Gly Leu Gly Gly Pro Asp Gly Lys Ser Gly Arg Trp Arg Arg
            200                 205                 210
Leu Gln Pro Arg Met Trp Ala Leu Phe Glu Asp Pro Tyr Ser Ser
        215                 220                 225
Arg Ala Ala Arg Phe Ile Ala Phe Ala Ser Leu Phe Phe Ile Leu
    230                 235                 240
Val Ser Ile Thr Thr Phe Cys Leu Glu Thr His Glu Ala Phe Asn
245                 250                 255
Ile Val Lys Asn Lys Thr Glu Pro Val Ile Asn Gly Thr Ser Val
            260                 265                 270
Val Leu Gln Tyr Glu Ile Glu Thr Asp Pro Ala Leu Thr Tyr Val
        275                 280                 285
Glu Gly Val Cys Val Val Trp Phe Thr Phe Glu Phe Leu Val Arg
    290                 295                 300
Ile Val Phe Ser Pro Asn Lys Leu Glu Phe Ile Lys Asn Leu Leu
305                 310                 315
Asn Ile Ile Asp Phe Val Ala Ile Leu Pro Phe Tyr Leu Glu Val
            320                 325                 330
Gly Leu Ser Gly Leu Ser Ser Lys Ala Ala Lys Asp Val Leu Gly
        335                 340                 345
Phe Leu Arg Val Val Arg Phe Val Arg Ile Leu Arg Ile Phe Lys
    350                 355                 360
Leu Thr Arg His Phe Val Gly Leu Arg Val Leu Gly His Thr Leu
365                 370                 375
Arg Ala Ser Thr Asn Glu Phe Leu Leu Leu Ile Ile Phe Leu Ala
            380                 385                 390
Leu Gly Val Leu Ile Phe Ala Thr Met Ile Tyr Tyr Ala Glu Arg
        395                 400                 405
Val Gly Ala Gln Pro Asn Asp Pro Ser Ala Ser Glu His Thr Gln
    410                 415                 420
Phe Lys Asn Ile Pro Ile Gly Phe Trp Trp Ala Val Val Thr Met
425                 430                 435
Thr Thr Leu Gly Tyr Gly Asp Met Tyr Pro Gln Thr Trp Ser Gly
            440                 445                 450
```

```
Met Leu Val Gly Ala Leu Cys Ala Leu Ala Gly Val Leu Thr Ile
                455                 460                 465
Ala Met Pro Val Pro Val Ile Val Asn Asn Phe Gly Met Tyr Tyr
                470                 475                 480
Ser Leu Ala Met Ala Lys Gln Lys Leu Pro Arg Lys Arg Lys Lys
                485                 490                 495
His Ile Pro Pro Ala Pro Gln Ala Ser Ser Pro Thr Phe Cys Lys
                500                 505                 510
Thr Glu Leu Asn Met Ala Cys Asn Ser Thr Gln Ser Asp Thr Cys
                515                 520                 525
Leu Gly Lys Asp Asn Arg Leu Leu Glu His Asn Arg Ser Val Leu
                530                 535                 540
Ser Gly Asp Asp Ser Thr Gly Ser Glu Pro Pro Leu Ser Pro Pro
                545                 550                 555
Glu Arg Leu Pro Ile Arg Arg Ser Ser Thr Arg Asp Lys Asn Arg
                560                 565                 570
Arg Gly Glu Thr Cys Phe Leu Leu Thr Thr Gly Asp Tyr Thr Cys
                575                 580                 585
Ala Ser Asp Gly Gly Ile Arg Lys Gly Tyr Glu Lys Ser Arg Ser
                590                 595                 600
Leu Asn Asn Ile Ala Gly Leu Ala Gly Asn Ala Leu Arg Leu Ser
                605                 610                 615
Pro Val Thr Ser Pro Tyr Asn Ser Pro Cys Pro Leu Arg Arg Ser
                620                 625                 630
Arg Ser Pro Ile Pro Ser Ile Leu
                635
```

<210> 19
<211> 681
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476949CD1

<400> 19

```
Met Ser Lys Asp Leu Ala Ala Met Gly Pro Gly Ala Ser Gly Asp
  1                 5                  10                  15
Gly Val Arg Thr Glu Thr Ala Pro His Ile Ala Leu Asp Ser Arg
                20                  25                  30
Val Gly Leu His Ala Tyr Asp Ile Ser Val Val Val Ile Tyr Phe
                35                  40                  45
Val Phe Val Ile Ala Val Gly Ile Trp Ser Ser Ile Arg Ala Ser
                50                  55                  60
Arg Gly Thr Ile Gly Gly Tyr Phe Leu Ala Gly Arg Ser Met Ser
                65                  70                  75
Trp Trp Pro Ile Gly Ala Ser Leu Met Ser Ser Asn Val Gly Ser
                80                  85                  90
Gly Leu Phe Ile Gly Leu Ala Gly Thr Gly Ala Ala Gly Gly Leu
                95                 100                 105
Ala Val Gly Gly Phe Glu Trp Asn Ala Thr Trp Leu Leu Leu Ala
               110                 115                 120
Leu Gly Trp Val Phe Val Pro Val Tyr Ile Ala Ala Gly Val Val
               125                 130                 135
Thr Met Pro Gln Tyr Leu Lys Lys Arg Phe Gly Gly Gln Arg Ile
               140                 145                 150
Gln Val Tyr Met Ser Val Leu Ser Leu Ile Leu Tyr Ile Phe Thr
               155                 160                 165
Lys Ile Ser Thr Asp Ile Phe Ser Gly Ala Leu Phe Ile Gln Met
               170                 175                 180
Ala Leu Gly Trp Asn Leu Tyr Leu Ser Thr Gly Ile Leu Leu Val
               185                 190                 195
Val Thr Ala Val Tyr Thr Ile Ala Gly Gly Leu Met Ala Val Ile
               200                 205                 210
Tyr Thr Asp Ala Leu Gln Thr Val Ile Met Val Gly Gly Ala Leu
               215                 220                 225
```

```
Val Leu Met Phe Leu Gly Phe Gln Asp Val Gly Trp Tyr Pro Gly
                230                 235                 240
Leu Glu Gln Arg Tyr Arg Gln Ala Ile Pro Asn Val Thr Val Pro
                245                 250                 255
Asn Thr Thr Cys His Leu Pro Arg Pro Asp Ala Phe His Ile Leu
                260                 265                 270
Arg Asp Pro Val Ser Gly Asp Ile Pro Trp Pro Gly Leu Ile Phe
                275                 280                 285
Gly Leu Thr Val Leu Ala Thr Trp Cys Trp Cys Thr Asp Gln Val
                290                 295                 300
Ile Val Gln Arg Ser Leu Ser Ala Lys Ser Leu Ser His Ala Lys
                305                 310                 315
Gly Gly Ser Val Leu Gly Gly Tyr Leu Lys Ile Leu Pro Met Phe
                320                 325                 330
Phe Ile Val Met Pro Gly Met Ile Ser Arg Ala Leu Phe Pro Asp
                335                 340                 345
Glu Val Gly Cys Val Asp Pro Asp Val Cys Gln Arg Ile Cys Gly
                350                 355                 360
Ala Arg Val Gly Cys Ser Asn Ile Ala Tyr Pro Lys Leu Val Met
                365                 370                 375
Ala Leu Met Pro Val Gly Leu Arg Gly Leu Met Ile Ala Val Ile
                380                 385                 390
Met Ala Ala Leu Met Ser Ser Leu Thr Ser Ile Phe Asn Ser Ser
                395                 400                 405
Ser Thr Leu Phe Thr Ile Asp Val Trp Gln Arg Phe Arg Arg Lys
                410                 415                 420
Ser Thr Glu Gln Glu Leu Met Val Val Gly Arg Val Phe Val Val
                425                 430                 435
Phe Leu Val Val Ile Ser Ile Leu Trp Ile Pro Ile Ile Gln Ser
                440                 445                 450
Ser Asn Ser Gly Gln Leu Phe Asp Tyr Ile Gln Ala Val Thr Ser
                455                 460                 465
Tyr Leu Ala Pro Pro Ile Thr Ala Leu Phe Leu Leu Ala Ile Phe
                470                 475                 480
Cys Lys Arg Val Thr Glu Pro Gly Ala Phe Trp Gly Leu Val Phe
                485                 490                 495
Gly Leu Gly Val Gly Leu Leu Arg Met Ile Leu Glu Phe Ser Tyr
                500                 505                 510
Pro Ala Pro Ala Cys Gly Glu Val Asp Arg Arg Pro Ala Val Leu
                515                 520                 525
Lys Asp Phe His Tyr Leu Tyr Phe Ala Ile Leu Leu Cys Gly Leu
                530                 535                 540
Thr Ala Ile Val Ile Val Ile Val Ser Leu Cys Thr Thr Pro Ile
                545                 550                 555
Pro Glu Glu Gln Leu Thr Arg Leu Thr Trp Trp Thr Arg Asn Cys
                560                 565                 570
Pro Leu Ser Glu Leu Glu Lys Glu Ala His Glu Ser Thr Pro Glu
                575                 580                 585
Ile Ser Glu Arg Pro Ala Gly Glu Cys Pro Ala Gly Gly Gly Ala
                590                 595                 600
Ala Glu Asn Ser Ser Leu Gly Gln Glu Gln Pro Glu Ala Pro Ser
                605                 610                 615
Arg Ser Trp Gly Lys Leu Leu Trp Ser Trp Phe Cys Gly Leu Ser
                620                 625                 630
Gly Thr Pro Glu Gln Ala Leu Ser Pro Ala Glu Lys Ala Ala Leu
                635                 640                 645
Glu Gln Lys Leu Thr Ser Ile Glu Glu Glu Pro Leu Trp Arg His
                650                 655                 660
Val Cys Asn Ile Asn Ala Val Leu Leu Leu Ala Ile Asn Ile Phe
                665                 670                 675
Leu Trp Gly Tyr Phe Ala
                680
```

<210> 20
<211> 1096
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477249CD1

<400> 20

```
Met Trp Arg Trp Ile Arg Gln Gln Leu Gly Phe Asp Pro Pro His
 1               5                  10                  15
Gln Ser Asp Thr Arg Thr Ile Tyr Val Ala Asn Arg Phe Pro Gln
            20                  25                  30
Asn Gly Leu Tyr Thr Pro Gln Lys Phe Ile Asp Asn Arg Ile Ile
        35                  40                  45
Ser Ser Lys Tyr Thr Val Trp Asn Phe Val Pro Lys Asn Leu Phe
    50                  55                  60
Glu Gln Phe Arg Arg Val Ala Asn Phe Tyr Phe Leu Ile Ile Phe
65                  70                  75
Leu Val Gln Leu Met Ile Asp Thr Pro Thr Ser Pro Val Thr Ser
                80                  85                  90
Gly Leu Pro Leu Phe Phe Val Ile Thr Val Thr Ala Ile Lys Gln
            95                  100                 105
Gly Tyr Glu Asp Trp Leu Arg His Asn Ser Asp Asn Glu Val Asn
        110                 115                 120
Gly Ala Pro Val Tyr Val Val Arg Ser Gly Gly Leu Val Lys Thr
    125                 130                 135
Arg Ser Lys Asn Ile Arg Val Gly Asp Ile Val Arg Ile Ala Lys
140                 145                 150
Asp Glu Ile Phe Pro Ala Asp Leu Val Leu Leu Ser Ser Asp Arg
                155                 160                 165
Leu Asp Gly Ser Cys His Val Thr Thr Ala Ser Leu Asp Gly Glu
            170                 175                 180
Thr Asn Leu Lys Thr His Val Ala Val Pro Glu Thr Ala Leu Leu
        185                 190                 195
Gln Thr Val Ala Asn Leu Asp Thr Leu Val Ala Val Ile Glu Cys
    200                 205                 210
Gln Gln Pro Glu Ala Asp Leu Tyr Arg Phe Met Gly Arg Met Ile
215                 220                 225
Ile Thr Gln Gln Met Glu Glu Ile Val Arg Pro Leu Gly Pro Glu
                230                 235                 240
Ser Leu Leu Leu Arg Gly Ala Arg Leu Lys Asn Thr Lys Glu Ile
            245                 250                 255
Phe Gly Val Ala Val Tyr Thr Gly Met Glu Thr Lys Met Ala Leu
        260                 265                 270
Asn Tyr Lys Ser Lys Ser Gln Lys Arg Ser Ala Val Glu Lys Ser
    275                 280                 285
Met Asn Thr Phe Leu Ile Ile Tyr Leu Val Ile Leu Ile Ser Glu
290                 295                 300
Ala Val Ile Ser Thr Ile Leu Lys Tyr Thr Trp Gln Ala Glu Glu
305                 310                 315
Lys Trp Asp Glu Pro Trp Tyr Asn Gln Lys Thr Glu His Gln Arg
                320                 325                 330
Asn Ser Ser Lys Val Glu Tyr Val Phe Thr Asp Lys Thr Gly Thr
            335                 340                 345
Leu Thr Glu Asn Glu Met Gln Phe Arg Glu Cys Ser Ile Asn Gly
        350                 355                 360
Met Lys Tyr Gln Glu Ile Asn Gly Arg Leu Val Pro Glu Gly Pro
    365                 370                 375
Thr Pro Asp Ser Ser Glu Gly Asn Leu Ser Tyr Leu Ser Ser Leu
380                 385                 390
Ser His Leu Asn Asn Leu Ser His Leu Thr Thr Ser Ser Ser Phe
395                 400                 405
Arg Thr Ser Pro Glu Asn Glu Thr Glu Leu Ile Lys Glu His Asp
                410                 415                 420
Leu Phe Phe Lys Ala Val Ser Leu Cys His Thr Val Gln Ile Ser
            425                 430                 435
Asn Val Gln Thr Asp Cys Thr Gly Asp Gly Pro Trp Gln Ser Asn
        440                 445                 450
Leu Ala Pro Ser Gln Leu Glu Tyr Tyr Ala Ser Ser Pro Asp Glu
    455                 460                 465
```

```
Lys Ala Leu Val Glu Ala Ala Ala Arg Ile Gly Ile Val Phe Ile
                470                 475                 480
Gly Asn Ser Glu Glu Thr Met Glu Val Lys Thr Leu Gly Lys Leu
            485                 490                 495
Glu Arg Tyr Lys Leu Leu His Ile Leu Glu Phe Asp Ser Asp Arg
        500                 505                 510
Arg Arg Met Ser Val Ile Val Gln Ala Pro Ser Gly Glu Lys Leu
    515                 520                 525
Leu Phe Ala Lys Gly Ala Glu Ser Ser Ile Leu Pro Lys Cys Ile
530                 535                 540
Gly Gly Glu Ile Glu Lys Thr Arg Ile His Val Asp Glu Phe Ala
545                 550                 555
Leu Lys Gly Leu Arg Thr Leu Cys Ile Ala Tyr Arg Lys Phe Thr
            560                 565                 570
Ser Lys Glu Tyr Glu Glu Ile Asp Lys Arg Ile Phe Glu Ala Arg
        575                 580                 585
Thr Ala Leu Gln Gln Arg Glu Glu Lys Leu Ala Ala Val Phe Gln
    590                 595                 600
Phe Ile Glu Lys Asp Leu Ile Leu Leu Gly Ala Thr Ala Val Glu
605                 610                 615
Asp Arg Leu Gln Asp Lys Val Arg Glu Thr Ile Glu Ala Leu Arg
            620                 625                 630
Met Ala Gly Ile Lys Val Trp Val Leu Thr Gly Asp Lys His Glu
        635                 640                 645
Thr Ala Val Ser Val Ser Leu Ser Cys Gly His Phe His Arg Thr
    650                 655                 660
Met Asn Ile Leu Glu Leu Ile Asn Gln Lys Ser Asp Ser Glu Cys
665                 670                 675
Ala Glu Gln Leu Arg Gln Leu Ala Arg Arg Ile Thr Glu Asp His
            680                 685                 690
Val Ile Gln His Gly Leu Val Val Asp Gly Thr Ser Leu Ser Leu
        695                 700                 705
Ala Leu Arg Glu His Glu Lys Leu Phe Met Glu Val Cys Arg Asn
    710                 715                 720
Cys Ser Ala Val Leu Cys Cys Arg Met Ala Pro Leu Gln Lys Ala
725                 730                 735
Lys Val Ile Arg Leu Ile Lys Ile Ser Pro Glu Lys Pro Ile Thr
            740                 745                 750
Leu Ala Val Gly Asp Gly Ala Asn Asp Val Ser Met Ile Gln Glu
        755                 760                 765
Ala His Val Gly Ile Gly Ile Met Gly Lys Glu Gly Arg Gln Ala
    770                 775                 780
Ala Arg Asn Ser Asp Tyr Ala Ile Ala Arg Phe Lys Phe Leu Ser
785                 790                 795
Lys Leu Leu Phe Val His Gly His Phe Tyr Tyr Ile Arg Ile Ala
            800                 805                 810
Thr Leu Val Gln Tyr Phe Phe Tyr Lys Asn Val Cys Phe Ile Thr
        815                 820                 825
Pro Gln Phe Leu Tyr Gln Phe Tyr Cys Leu Phe Ser Gln Gln Thr
    830                 835                 840
Leu Tyr Asp Ser Val Tyr Leu Thr Leu Tyr Asn Ile Cys Phe Thr
845                 850                 855
Ser Leu Pro Ile Leu Ile Tyr Ser Leu Leu Glu Gln His Val Asp
            860                 865                 870
Pro His Val Leu Gln Asn Lys Pro Thr Leu Tyr Arg Asp Ile Ser
        875                 880                 885
Lys Asn Arg Leu Leu Ser Ile Lys Thr Phe Leu Tyr Trp Thr Ile
    890                 895                 900
Leu Gly Phe Ser His Ala Phe Ile Phe Phe Phe Gly Ser Tyr Leu
905                 910                 915
Leu Ile Gly Lys Asp Thr Ser Leu Leu Gly Asn Gly Gln Met Phe
            920                 925                 930
Gly Asn Trp Thr Phe Gly Thr Leu Val Phe Thr Val Met Val Ile
        935                 940                 945
Thr Val Thr Val Lys Met Ala Leu Glu Thr His Phe Trp Thr Trp
    950                 955                 960
Ile Asn His Leu Val Thr Trp Gly Ser Ile Ile Phe Tyr Phe Val
```

965 970 975
Phe Ser Leu Phe Tyr Gly Gly Ile Leu Trp Pro Phe Leu Gly Ser
980 985 990
Gln Asn Met Tyr Phe Val Phe Ile Gln Leu Leu Ser Ser Gly Ser
995 1000 1005
Ala Trp Phe Ala Ile Ile Leu Met Val Val Thr Cys Leu Phe Leu
1010 1015 1020
Asp Ile Ile Lys Lys Val Phe Asp Arg His Leu His Pro Thr Ser
1025 1030 1035
Thr Glu Lys Ala Gln Leu Thr Glu Thr Asn Ala Gly Ile Lys Cys
1040 1045 1050
Leu Asp Ser Met Cys Cys Phe Pro Glu Gly Glu Ala Ala Cys Ala
1055 1060 1065
Ser Val Gly Arg Met Leu Glu Arg Val Ile Gly Arg Cys Ser Pro
1070 1075 1080
Thr His Ile Ser Arg Cys Glu Ile Ser Leu Ser Ser Leu Cys Cys
1085 1090 1095
Arg

<210> 21
<211> 707
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477720CD1

<400> 21
Met Ala Leu Gln Met Phe Val Thr Tyr Ser Pro Trp Asn Cys Leu
1 5 10 15
Leu Leu Leu Val Ala Leu Glu Cys Ser Glu Ala Ser Ser Asp Leu
20 25 30
Asn Glu Ser Ala Asn Ser Thr Ala Gln Tyr Ala Ser Asn Ala Trp
35 40 45
Phe Ala Ala Ala Ser Ser Glu Pro Glu Glu Gly Ile Ser Val Phe
50 55 60
Glu Leu Asp Tyr Asp Tyr Val Gln Ile Pro Tyr Glu Val Thr Leu
65 70 75
Trp Ile Leu Leu Ala Ser Leu Ala Lys Ile Gly Phe His Leu Tyr
80 85 90
His Arg Leu Pro Gly Leu Met Pro Glu Ser Cys Leu Leu Ile Leu
95 100 105
Val Gly Ala Leu Val Gly Gly Ile Ile Phe Gly Thr Asp His Lys
110 115 120
Ser Pro Pro Val Met Asp Ser Ser Ile Tyr Phe Leu Tyr Leu Leu
125 130 135
Pro Pro Ile Val Leu Glu Gly Gly Tyr Phe Met Pro Thr Arg Pro
140 145 150
Phe Phe Glu Asn Ile Gly Ser Ile Leu Trp Trp Ala Val Leu Gly
155 160 165
Ala Leu Ile Asn Ala Leu Gly Ile Gly Leu Ser Leu Tyr Leu Ile
170 175 180
Cys Gln Val Lys Ala Phe Gly Leu Gly Asp Val Asn Leu Leu Gln
185 190 195
Asn Leu Leu Phe Gly Ser Leu Ile Ser Ala Val Asp Pro Val Ala
200 205 210
Val Leu Ala Val Phe Glu Glu Ala Arg Val Asn Glu Gln Leu Tyr
215 220 225
Met Met Ile Phe Gly Glu Ala Leu Leu Asn Asp Gly Ile Thr Val
230 235 240
Val Leu Tyr Asn Met Leu Ile Ala Phe Thr Lys Met His Lys Phe
245 250 255
Glu Asp Ile Glu Thr Val Asp Ile Leu Ala Gly Cys Ala Arg Phe
260 265 270
Ile Val Val Gly Leu Gly Gly Val Leu Phe Gly Ile Val Phe Gly

275 280 285
Phe Ile Ser Ala Phe Ile Thr Arg Phe Thr Gln Asn Ile Ser Ala
290 295 300
Ile Glu Pro Leu Ile Val Phe Met Phe Ser Tyr Leu Ser Tyr Leu
305 310 315
Ala Ala Glu Thr Leu Tyr Leu Ser Gly Ile Leu Ala Ile Thr Ala
320 325 330
Cys Ala Val Thr Met Lys Lys Tyr Val Glu Glu Asn Val Ser Gln
335 340 345
Thr Ser Tyr Thr Thr Ile Lys Tyr Phe Met Lys Met Leu Ser Ser
350 355 360
Val Ser Glu Thr Leu Ile Phe Ile Phe Met Gly Val Ser Thr Val
365 370 375
Gly Lys Asn His Glu Trp Asn Trp Ala Phe Ile Cys Phe Thr Leu
380 385 390
Ala Phe Cys Gln Ile Trp Arg Ala Ile Ser Val Phe Ala Leu Phe
395 400 405
Tyr Ile Ser Asn Gln Phe Arg Thr Phe Pro Phe Ser Ile Lys Asp
410 415 420
Gln Cys Ile Ile Phe Tyr Ser Gly Val Arg Gly Ala Gly Ser Phe
425 430 435
Ser Leu Ala Phe Leu Leu Pro Leu Ser Leu Phe Pro Arg Lys Lys
440 445 450
Met Phe Val Thr Ala Thr Leu Val Val Ile Tyr Phe Thr Val Phe
455 460 465
Ile Gln Gly Ile Thr Val Gly Pro Leu Val Arg Tyr Leu Asp Val
470 475 480
Lys Lys Thr Asn Lys Lys Glu Ser Ile Asn Glu Glu Leu His Ile
485 490 495
Arg Leu Met Asp His Leu Lys Ala Gly Ile Glu Asp Val Cys Gly
500 505 510
His Trp Ser His Tyr Gln Val Arg Asp Lys Phe Lys Lys Phe Asp
515 520 525
His Arg Tyr Leu Arg Lys Ile Leu Ile Arg Lys Asn Leu Pro Lys
530 535 540
Ser Ser Ile Val Ser Leu Tyr Lys Lys Leu Glu Met Lys Gln Ala
545 550 555
Ile Glu Met Val Glu Thr Gly Ile Leu Ser Ser Thr Ala Phe Ser
560 565 570
Ile Pro His Gln Ala Gln Arg Ile Gln Gly Ile Lys Arg Leu Ser
575 580 585
Pro Glu Asp Val Glu Ser Ile Arg Asp Ile Leu Thr Ser Asn Met
590 595 600
Tyr Gln Val Arg Gln Arg Thr Leu Ser Tyr Asn Lys Tyr Asn Leu
605 610 615
Lys Pro Gln Thr Ser Glu Lys Gln Ala Lys Glu Ile Leu Ile Arg
620 625 630
Arg Gln Asn Thr Leu Arg Glu Ser Met Arg Lys Gly His Ser Leu
635 640 645
Pro Trp Gly Lys Pro Ala Gly Thr Lys Asn Ile Arg Tyr Leu Ser
650 655 660
Tyr Pro Tyr Gly Asn Pro Gln Ser Ala Gly Arg Asp Thr Arg Ala
665 670 675
Ala Gly Phe Ser Gly Lys Leu Pro Thr Trp Leu Leu Cys Cys Phe
680 685 690
Ser Val Glu Ser Gly Gly Lys Tyr Leu Gly Val Trp Ala Lys Arg
695 700 705
Gln His

<210> 22
<211> 729
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7477852CD1

<400> 22

```
Met Gly Gly Phe Leu Pro Lys Ala Glu Gly Pro Gly Ser Gln Leu
  1               5                  10                  15
Gln Lys Leu Leu Pro Ser Phe Leu Val Arg Glu Gln Asp Trp Asp
             20                  25                  30
Gln His Leu Asp Lys Leu His Met Leu Gln Gln Lys Arg Ile Leu
         35                  40                  45
Glu Ser Pro Leu Leu Arg Ala Ser Lys Glu Asn Asp Leu Ser Val
     50                  55                  60
Leu Arg Gln Leu Leu Leu Asp Cys Thr Cys Asp Val Arg Gln Arg
 65                  70                  75
Gly Ala Leu Gly Glu Thr Ala Leu His Ile Ala Ala Leu Tyr Asp
             80                  85                  90
Asn Leu Glu Ala Ala Leu Val Leu Met Glu Ala Ala Pro Glu Leu
         95                 100                 105
Val Phe Glu Pro Thr Thr Cys Glu Ala Phe Ala Gly Gln Thr Ala
    110                 115                 120
Leu His Ile Ala Val Val Asn Gln Asn Val Asn Leu Val Arg Ala
125                 130                 135
Leu Leu Thr Arg Arg Ala Ser Val Ser Ala Arg Ala Thr Gly Thr
            140                 145                 150
Ala Phe Arg Arg Ser Pro Arg Asn Leu Ile Tyr Phe Gly Glu His
        155                 160                 165
Pro Leu Ser Phe Ala Ala Cys Val Asn Ser Glu Glu Ile Val Arg
    170                 175                 180
Leu Leu Ile Glu His Gly Ala Asp Ile Arg Ala Gln Asp Ser Leu
185                 190                 195
Gly Asn Thr Val Leu His Ile Leu Ile Leu Gln Pro Asn Lys Thr
            200                 205                 210
Phe Ala Cys Gln Met Tyr Asn Leu Leu Leu Ser Tyr Asp Gly His
        215                 220                 225
Gly Asp His Leu Gln Pro Leu Asp Leu Val Pro Asn His Gln Gly
    230                 235                 240
Leu Thr Pro Phe Lys Leu Ala Gly Val Glu Gly Asn Thr Val Met
245                 250                 255
Phe Gln His Leu Met Gln Lys Arg Arg His Ile Gln Trp Thr Tyr
            260                 265                 270
Gly Pro Leu Thr Ser Ile Leu Tyr Asp Leu Thr Glu Ile Asp Ser
        275                 280                 285
Trp Gly Glu Glu Leu Ser Phe Leu Glu Leu Val Val Ser Ser Asp
    290                 295                 300
Lys Arg Glu Ala Arg Gln Ile Leu Glu Gln Thr Pro Val Lys Glu
305                 310                 315
Leu Val Ser Phe Lys Trp Asn Lys Tyr Gly Arg Pro Tyr Phe Cys
            320                 325                 330
Ile Leu Ala Ala Leu Tyr Leu Leu Tyr Met Ile Cys Phe Thr Thr
        335                 340                 345
Cys Cys Val Tyr Arg Pro Leu Lys Phe Arg Gly Gly Asn Arg Thr
    350                 355                 360
His Ser Arg Asp Ile Thr Ile Leu Gln Gln Lys Leu Leu Gln Glu
365                 370                 375
Ala Tyr Glu Thr Arg Glu Asp Ile Ile Arg Leu Val Gly Glu Leu
            380                 385                 390
Val Ser Ile Val Gly Ala Val Ile Ile Leu Leu Leu Glu Ile Pro
        395                 400                 405
Asp Ile Phe Arg Val Gly Ala Ser Arg Tyr Phe Gly Lys Thr Ile
    410                 415                 420
Leu Gly Gly Pro Phe His Val Ile Met Ile Thr Tyr Ala Ser Leu
425                 430                 435
Val Leu Val Thr Met Val Met Arg Leu Thr Asn Thr Asn Gly Glu
            440                 445                 450
Val Val Pro Met Ser Phe Ala Leu Val Leu Gly Trp Cys Ser Val
        455                 460                 465
Met Tyr Phe Thr Arg Gly Phe Gln Met Leu Gly Pro Phe Thr Ile
    470                 475                 480
```

```
Met Ile Gln Lys Met Ile Phe Gly Asp Leu Met Arg Phe Cys Trp
                485                 490                 495
Leu Met Ala Val Val Ile Leu Gly Phe Ala Ser Ala Phe Tyr Ile
            500                 505                 510
Ile Phe Gln Thr Glu Asp Pro Thr Ser Leu Gly Gln Phe Tyr Asp
        515                 520                 525
Tyr Pro Met Ala Leu Phe Thr Thr Phe Glu Leu Phe Leu Thr Val
    530                 535                 540
Ile Asp Ala Pro Ala Asn Tyr Asp Val Asp Leu Pro Phe Met Phe
545                 550                 555
Ser Ile Val Asn Phe Ala Phe Ala Ile Ile Ala Thr Leu Leu Met
                560                 565                 570
Leu Asn Leu Phe Ile Ala Met Met Gly Asp Thr His Trp Arg Val
            575                 580                 585
Ala Gln Glu Arg Asp Glu Leu Trp Arg Ala Gln Val Val Ala Thr
        590                 595                 600
Thr Val Met Leu Glu Arg Lys Leu Pro Arg Cys Leu Trp Pro Arg
    605                 610                 615
Ser Gly Ile Cys Gly Cys Glu Phe Gly Leu Gly Asp Arg Trp Phe
620                 625                 630
Leu Arg Val Glu Asn His Asn Asp Gln Asn Pro Leu Arg Val Leu
                635                 640                 645
Arg Tyr Val Glu Val Phe Lys Asn Ser Asp Lys Glu Asp Asp Gln
            650                 655                 660
Glu His Pro Ser Glu Lys Gln Pro Ser Gly Ala Glu Ser Gly Thr
        665                 670                 675
Leu Ala Arg Ala Ser Leu Ala Leu Pro Thr Ser Ser Leu Ser Arg
    680                 685                 690
Thr Ala Ser Gln Ser Ser Ser His Arg Gly Trp Glu Ile Leu Arg
695                 700                 705
Gln Asn Thr Leu Gly His Leu Asn Leu Gly Leu Asn Leu Ser Glu
                710                 715                 720
Gly Asp Gly Glu Glu Val Tyr His Phe
            725
```

<210> 23
<211> 492
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1471717CD1

<400> 23

```
Met Ala Thr Lys Pro Thr Glu Pro Val Thr Ile Leu Ser Leu Arg
  1                 5                  10                  15
Lys Leu Ser Leu Gly Thr Ala Glu Pro Gln Val Lys Glu Pro Lys
            20                  25                  30
Thr Phe Thr Val Glu Asp Ala Val Glu Thr Ile Gly Phe Gly Arg
        35                  40                  45
Phe His Ile Ala Leu Phe Leu Ile Met Gly Ser Thr Gly Val Val
    50                  55                  60
Glu Ala Met Glu Ile Met Leu Ile Ala Val Val Ser Pro Val Ile
65                  70                  75
Arg Cys Glu Trp Gln Leu Glu Asn Trp Gln Val Ala Leu Val Thr
                80                  85                  90
Thr Met Val Phe Phe Gly Tyr Met Val Phe Ser Ile Leu Phe Gly
            95                  100                 105
Leu Leu Ala Asp Arg Tyr Gly Arg Trp Lys Ile Leu Leu Ile Ser
        110                 115                 120
Phe Leu Trp Gly Ala Tyr Phe Ser Leu Leu Thr Ser Phe Ala Pro
    125                 130                 135
Ser Tyr Ile Trp Phe Val Phe Leu Arg Thr Met Val Gly Cys Gly
140                 145                 150
Val Ser Gly His Ser Gln Gly Leu Ile Ile Lys Thr Glu Phe Leu
                155                 160                 165
```

```
Pro Thr Lys Tyr Arg Gly Tyr Met Leu Pro Leu Ser Gln Val Phe
                170                 175                 180
Trp Leu Ala Gly Ser Leu Leu Ile Ile Gly Leu Ala Ser Val Ile
                185                 190                 195
Ile Pro Thr Ile Gly Trp Arg Trp Leu Ile Arg Val Ala Ser Ile
                200                 205                 210
Pro Gly Ile Ile Leu Ile Val Ala Phe Lys Phe Ile Pro Glu Ser
                215                 220                 225
Ala Arg Phe Asn Val Ser Thr Gly Asn Thr Arg Ala Ala Leu Ala
                230                 235                 240
Thr Leu Glu Arg Val Ala Lys Met Asn Arg Ser Val Met Pro Glu
                245                 250                 255
Gly Lys Leu Val Glu Pro Val Leu Glu Lys Arg Gly Arg Phe Ala
                260                 265                 270
Asp Leu Leu Asp Ala Lys Tyr Leu Arg Thr Thr Leu Gln Ile Trp
                275                 280                 285
Val Ile Trp Leu Gly Ile Ser Phe Ala Tyr Tyr Gly Val Ile Leu
                290                 295                 300
Ala Ser Ala Glu Leu Leu Glu Arg Asp Leu Val Cys Gly Ser Lys
                305                 310                 315
Ser Asp Ser Ala Val Val Val Thr Gly Gly Asp Ser Gly Glu Ser
                320                 325                 330
Gln Ser Pro Cys Tyr Cys His Met Phe Ala Pro Ser Asp Tyr Arg
                335                 340                 345
Thr Met Ile Ile Ser Thr Ile Gly Glu Ile Ala Leu Asn Pro Leu
                350                 355                 360
Asn Ile Leu Gly Ile Asn Phe Leu Gly Arg Arg Leu Ser Leu Ser
                365                 370                 375
Ile Thr Met Gly Cys Thr Ala Leu Phe Cys Leu Leu Leu Asn Ile
                380                 385                 390
Cys Thr Ser Ser Ala Gly Leu Ile Gly Phe Leu Phe Met Leu Arg
                395                 400                 405
Ala Leu Val Ala Ala Asn Phe Asn Thr Val Tyr Ile Tyr Thr Ala
                410                 415                 420
Glu Val Tyr Pro Thr Thr Met Arg Ala Leu Gly Met Gly Thr Ser
                425                 430                 435
Gly Ser Leu Cys Arg Ile Gly Ala Met Val Ala Pro Phe Ile Ser
                440                 445                 450
Gln Val Leu Met Ser Ala Ser Ile Leu Gly Ala Leu Cys Leu Phe
                455                 460                 465
Ser Ser Val Cys Val Val Cys Ala Ile Ser Ala Phe Thr Leu Pro
                470                 475                 480
Ile Glu Thr Lys Gly Arg Ala Leu Gln Gln Ile Lys
                485                 490
```

<210> 24
<211> 1494
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3874406CD1

<400> 24

```
Met Asn Met Lys Gln Lys Ser Val Tyr Gln Gln Thr Lys Ala Leu
  1               5                  10                  15
Leu Cys Lys Asn Phe Leu Lys Lys Trp Arg Met Lys Arg Glu Ser
                20                  25                  30
Leu Leu Glu Trp Gly Leu Ser Ile Leu Leu Gly Leu Cys Ile Ala
                35                  40                  45
Leu Phe Ser Ser Ser Met Arg Asn Val Gln Phe Pro Gly Met Ala
                50                  55                  60
Pro Gln Asn Leu Gly Arg Val Asp Lys Phe Asn Ser Ser Ser Leu
                65                  70                  75
Met Val Val Tyr Thr Pro Ile Ser Asn Leu Thr Gln Gln Ile Met
                80                  85                  90
```

```
Asn Lys Thr Ala Leu Ala Pro Leu Leu Lys Gly Thr Ser Val Ile
 95                 100                 105
Gly Ala Pro Asn Lys Thr His Met Asp Glu Ile Leu Leu Glu Asn
110                 115                 120
Leu Pro Tyr Ala Met Gly Ile Ile Phe Asn Glu Thr Phe Ser Tyr
125                 130                 135
Lys Leu Ile Phe Phe Gln Gly Tyr Asn Ser Pro Leu Trp Lys Glu
140                 145                 150
Asp Phe Ser Ala His Cys Trp Asp Gly Tyr Gly Glu Phe Ser Cys
155                 160                 165
Thr Leu Thr Lys Tyr Trp Asn Arg Gly Phe Val Ala Leu Gln Thr
170                 175                 180
Ala Ile Asn Thr Ala Ile Ile Glu Val Ala Leu Val Phe Leu Met
185                 190                 195
Ser Val Leu Leu Lys Lys Ala Val Leu Thr Asn Leu Val Val Phe
200                 205                 210
Leu Leu Thr Leu Phe Trp Gly Cys Leu Gly Phe Thr Val Phe Tyr
215                 220                 225
Glu Gln Leu Pro Ser Ser Leu Glu Trp Ile Leu Asn Ile Cys Ser
230                 235                 240
Pro Phe Ala Phe Thr Thr Gly Met Ile Gln Ile Ile Lys Leu Asp
245                 250                 255
Tyr Asn Leu Asn Gly Val Ile Phe Pro Asp Pro Ser Gly Asp Ser
260                 265                 270
Tyr Thr Met Ile Ala Thr Phe Ser Met Leu Leu Leu Asp Gly Leu
275                 280                 285
Ile Tyr Leu Leu Leu Ala Leu Tyr Phe Asp Lys Ile Leu Pro Tyr
290                 295                 300
Gly Asp Glu Arg His Tyr Ser Pro Leu Phe Phe Leu Asn Ser Ser
305                 310                 315
Ser Cys Phe Gln His Gln Arg Thr Asn Ala Lys Val Ile Glu Lys
320                 325                 330
Glu Ile Asp Ala Glu His Pro Ser Asp Asp Tyr Phe Glu Pro Val
335                 340                 345
Ala Pro Glu Phe Gln Gly Lys Glu Ala Ile Arg Ile Arg Asn Val
350                 355                 360
Lys Lys Glu Tyr Lys Gly Lys Ser Gly Lys Val Glu Ala Leu Lys
365                 370                 375
Gly Leu Leu Phe Asp Ile Tyr Glu Gly Gln Ile Thr Ala Ile Leu
380                 385                 390
Gly His Ser Gly Ala Gly Lys Ser Ser Leu Leu Asn Ile Leu Asn
395                 400                 405
Gly Leu Ser Val Pro Thr Glu Gly Ser Val Thr Ile Tyr Asn Lys
410                 415                 420
Asn Leu Ser Glu Met Gln Asp Leu Glu Glu Ile Arg Lys Ile Thr
425                 430                 435
Gly Val Cys Pro Gln Phe Asn Val Gln Phe Asp Ile Leu Thr Val
440                 445                 450
Lys Glu Asn Leu Ser Leu Phe Ala Lys Ile Lys Gly Ile His Leu
455                 460                 465
Lys Glu Val Glu Gln Glu Val Gln Arg Ile Leu Leu Glu Leu Asp
470                 475                 480
Met Gln Asn Ile Gln Asp Asn Leu Ala Lys His Leu Ser Glu Gly
485                 490                 495
Gln Lys Arg Lys Leu Thr Phe Gly Ile Thr Ile Leu Gly Asp Pro
500                 505                 510
Gln Ile Leu Leu Leu Asp Glu Pro Thr Thr Gly Leu Asp Pro Phe
515                 520                 525
Ser Arg Asp Gln Val Trp Ser Leu Leu Arg Glu Arg Arg Ala Asp
530                 535                 540
His Val Ile Leu Phe Ser Thr Gln Ser Met Asp Glu Ala Asp Ile
545                 550                 555
Leu Ala Asp Arg Lys Val Ile Met Ser Asn Gly Arg Leu Lys Cys
560                 565                 570
Ala Gly Ser Ser Ile Phe Leu Lys Arg Arg Trp Gly Leu Gly Tyr
575                 580                 585
His Leu Ser Leu His Arg Asn Glu Ile Cys Asn Pro Glu Gln Ile
```

590 595 600
Thr Ser Phe Ile Thr His His Ile Pro Asp Ala Lys Leu Lys Thr
605 610 615
Glu Asn Lys Glu Lys Leu Val Tyr Thr Leu Pro Leu Glu Arg Thr
620 625 630
Asn Thr Phe Pro Asp Leu Phe Ser Asp Leu Asp Lys Cys Ser Asp
635 640 645
Gln Gly Val Thr Gly Tyr Asp Ile Ser Met Ser Thr Leu Asn Glu
650 655 660
Val Phe Met Lys Leu Glu Gly Gln Ser Thr Ile Glu Gln Asp Phe
665 670 675
Glu Gln Val Glu Met Ile Arg Asp Ser Glu Ser Leu Asn Glu Met
680 685 690
Glu Leu Ala His Ser Ser Phe Ser Glu Met Gln Thr Ala Val Ser
695 700 705
Asp Met Gly Leu Trp Arg Met Gln Val Phe Ala Met Ala Arg Leu
710 715 720
Arg Phe Leu Lys Leu Lys Arg Gln Thr Lys Val Leu Leu Thr Leu
725 730 735
Leu Leu Val Phe Gly Ile Ala Ile Phe Pro Leu Ile Val Glu Asn
740 745 750
Ile Ile Tyr Ala Met Leu Asn Glu Lys Ile Asp Trp Glu Phe Lys
755 760 765
Asn Glu Leu Tyr Phe Leu Ser Pro Gly Gln Leu Pro Gln Glu Pro
770 775 780
Arg Thr Ser Leu Leu Ile Ile Asn Asn Thr Glu Ser Asn Ile Glu
785 790 795
Asp Phe Ile Lys Ser Leu Lys His Gln Asn Ile Leu Leu Glu Val
800 805 810
Asp Asp Phe Glu Asn Arg Asn Gly Thr Asp Gly Leu Ser Tyr Asn
815 820 825
Gly Ala Ile Ile Val Ser Gly Lys Gln Lys Asp Tyr Arg Phe Ser
830 835 840
Val Val Cys Asn Thr Lys Arg Leu His Cys Phe Pro Ile Leu Met
845 850 855
Asn Ile Ile Ser Asn Gly Leu Leu Gln Met Phe Asn His Thr Gln
860 865 870
His Ile Arg Ile Glu Ser Ser Pro Phe Pro Leu Ser His Ile Gly
875 880 885
Leu Trp Thr Gly Leu Pro Asp Gly Ser Phe Phe Leu Phe Leu Val
890 895 900
Leu Cys Ser Ile Ser Pro Tyr Ile Thr Met Gly Ser Ile Ser Asp
905 910 915
Tyr Lys Lys Asn Ala Lys Ser Gln Leu Trp Ile Ser Gly Leu Tyr
920 925 930
Thr Ser Ala Tyr Trp Cys Gly Gln Ala Leu Val Asp Val Ser Phe
935 940 945
Phe Ile Leu Ile Leu Leu Leu Met Tyr Leu Ile Phe Tyr Ile Glu
950 955 960
Asn Met Gln Tyr Leu Leu Ile Thr Ser Gln Ile Val Phe Ala Leu
965 970 975
Val Ile Val Thr Pro Gly Tyr Ala Ala Ser Leu Val Phe Phe Ile
980 985 990
Tyr Met Ile Ser Phe Ile Phe Arg Lys Arg Arg Lys Asn Ser Gly
995 1000 1005
Leu Trp Ser Phe Tyr Phe Phe Phe Ala Ser Thr Ile Met Phe Ser
1010 1015 1020
Ile Thr Leu Ile Asn His Phe Asp Leu Ser Ile Leu Ile Thr Thr
1025 1030 1035
Met Val Leu Val Pro Ser Tyr Thr Leu Leu Gly Phe Lys Thr Phe
1040 1045 1050
Leu Glu Val Arg Asp Gln Glu His Tyr Arg Glu Phe Pro Glu Ala
1055 1060 1065
Asn Phe Glu Leu Ser Ala Thr Asp Phe Leu Val Cys Phe Ile Pro
1070 1075 1080
Tyr Phe Gln Thr Leu Leu Phe Val Phe Val Leu Arg Cys Met Glu
1085 1090 1095

```
Leu Lys Cys Gly Lys Lys Arg Met Arg Lys Asp Pro Val Phe Arg
                1100                1105                1110
Ile Ser Pro Gln Ser Arg Asp Ala Lys Pro Asn Pro Glu Glu Pro
                1115                1120                1125
Ile Asp Glu Asp Glu Asp Ile Gln Thr Glu Arg Ile Arg Thr Val
                1130                1135                1140
Thr Ala Leu Thr Thr Ser Ile Leu Asp Glu Lys Pro Val Ile Ile
                1145                1150                1155
Ala Ser Cys Leu His Lys Glu Tyr Ala Gly Gln Lys Lys Ser Cys
                1160                1165                1170
Phe Ser Lys Arg Lys Lys Lys Ile Ala Ala Arg Asn Ile Ser Phe
                1175                1180                1185
Cys Val Gln Glu Gly Glu Ile Leu Gly Leu Leu Gly Pro Ser Gly
                1190                1195                1200
Ala Gly Lys Ser Ser Ser Ile Arg Met Ile Ser Gly Ile Thr Lys
                1205                1210                1215
Pro Thr Ala Gly Glu Val Glu Leu Lys Gly Cys Ser Ser Val Leu
                1220                1225                1230
Gly His Leu Gly Tyr Cys Pro Gln Glu Asn Val Leu Trp Pro Met
                1235                1240                1245
Leu Thr Leu Arg Glu His Leu Glu Val Tyr Ala Ala Val Lys Gly
                1250                1255                1260
Leu Arg Glu Ala Asp Ala Arg Leu Ala Ile Ala Arg Leu Val Ser
                1265                1270                1275
Ala Phe Lys Leu His Glu Gln Leu Asn Val Pro Val Gln Lys Leu
                1280                1285                1290
Thr Ala Gly Ile Thr Arg Lys Leu Cys Phe Val Leu Ser Leu Leu
                1295                1300                1305
Gly Asn Ser Pro Val Leu Leu Leu Asp Glu Pro Ser Thr Gly Ile
                1310                1315                1320
Asp Pro Thr Gly Gln Gln Gln Met Trp Gln Ala Ile Gln Ala Val
                1325                1330                1335
Val Lys Asn Thr Glu Arg Gly Val Leu Leu Thr Thr His Asn Leu
                1340                1345                1350
Ala Glu Ala Glu Ala Leu Cys Asp Arg Val Ala Ile Met Val Ser
                1355                1360                1365
Gly Arg Leu Arg Cys Ile Gly Ser Ile Gln His Leu Lys Asn Lys
                1370                1375                1380
Leu Gly Lys Asp Tyr Ile Leu Glu Leu Lys Val Lys Glu Thr Ser
                1385                1390                1395
Gln Val Thr Leu Val His Thr Glu Ile Leu Lys Leu Phe Pro Gln
                1400                1405                1410
Ala Ala Gly Gln Gln Arg Tyr Ser Ser Leu Leu Thr Tyr Lys Leu
                1415                1420                1425
Pro Val Ala Asp Val Tyr Pro Leu Ser Gln Thr Phe His Lys Leu
                1430                1435                1440
Glu Ala Val Lys His Asn Phe Asn Leu Glu Glu Tyr Ser Leu Ser
                1445                1450                1455
Gln Cys Thr Leu Glu Lys Val Phe Leu Glu Leu Ser Lys Glu Gln
                1460                1465                1470
Glu Val Gly Asn Phe Asp Glu Glu Ile Asp Thr Thr Met Arg Trp
                1475                1480                1485
Lys Leu Leu Pro His Ser Asp Glu Pro
                1490
```

<210> 25
<211> 774
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4599654CD1

<400> 25

```
Met Glu Ala Glu Gln Arg Pro Ala Ala Gly Ala Ser Glu Gly Ala
  1               5                  10                  15
```

```
Thr Pro Gly Leu Glu Ala Val Pro Pro Val Ala Pro Pro Pro Ala
                20                  25                  30
Thr Ala Ala Ser Gly Pro Ile Pro Lys Ser Gly Pro Glu Pro Lys
            35                  40                  45
Arg Arg His Leu Gly Thr Leu Leu Gln Pro Thr Val Asn Lys Phe
        50                  55                  60
Ser Leu Arg Val Phe Gly Ser His Lys Ala Val Glu Ile Glu Gln
    65                  70                  75
Glu Arg Val Lys Ser Ala Gly Ala Trp Ile Ile His Pro Tyr Ser
                80                  85                  90
Asp Phe Arg Phe Tyr Trp Asp Leu Ile Met Leu Leu Leu Met Val
            95                  100                 105
Gly Asn Leu Ile Val Leu Pro Val Gly Ile Thr Phe Phe Lys Glu
        110                 115                 120
Glu Asn Ser Pro Pro Trp Ile Val Phe Asn Val Leu Ser Asp Thr
    125                 130                 135
Phe Phe Leu Leu Asp Leu Val Leu Asn Phe Arg Thr Gly Ile Val
140                 145                 150
Val Glu Glu Gly Ala Glu Ile Leu Leu Ala Pro Arg Ala Ile Arg
155                 160                 165
Thr Arg Tyr Leu Arg Thr Trp Phe Leu Val Asp Leu Ile Ser Ser
170                 175                 180
Ile Pro Val Asp Tyr Ile Phe Leu Val Val Glu Leu Glu Pro Arg
185                 190                 195
Leu Asp Ala Glu Val Tyr Lys Thr Ala Arg Ala Leu Arg Ile Val
200                 205                 210
Arg Phe Thr Lys Ile Leu Ser Leu Leu Arg Leu Leu Arg Leu Ser
215                 220                 225
Arg Leu Ile Arg Tyr Ile His Gln Trp Glu Glu Ile Phe His Met
230                 235                 240
Thr Tyr Asp Leu Ala Ser Ala Val Val Arg Ile Phe Asn Leu Ile
245                 250                 255
Gly Met Met Leu Leu Leu Cys His Trp Asp Gly Cys Leu Gln Phe
260                 265                 270
Leu Val Pro Met Leu Gln Asp Phe Pro Pro Asp Cys Trp Val Ser
275                 280                 285
Ile Asn His Met Val Asn His Ser Trp Gly Arg Gln Tyr Ser His
290                 295                 300
Ala Leu Phe Lys Ala Met Ser His Met Leu Cys Ile Gly Tyr Gly
305                 310                 315
Gln Gln Ala Pro Val Gly Met Pro Asp Val Trp Leu Thr Met Leu
320                 325                 330
Ser Met Ile Val Gly Ala Thr Cys Tyr Ala Met Phe Ile Gly His
335                 340                 345
Ala Thr Ala Leu Ile Gln Ser Leu Asp Ser Ser Arg Arg Gln Tyr
350                 355                 360
Gln Glu Lys Tyr Lys Gln Val Glu Gln Tyr Met Ser Phe His Lys
365                 370                 375
Leu Pro Ala Asp Thr Arg Gln Arg Ile His Glu Tyr Tyr Glu His
380                 385                 390
Arg Tyr Gln Gly Lys Met Phe Asp Glu Glu Ser Ile Leu Gly Glu
395                 400                 405
Leu Ser Glu Pro Leu Arg Glu Glu Ile Ile Asn Phe Thr Cys Arg
410                 415                 420
Gly Leu Val Ala His Met Pro Leu Phe Ala His Ala Asp Pro Ser
425                 430                 435
Phe Val Thr Ala Val Leu Thr Lys Leu Arg Phe Glu Val Phe Gln
440                 445                 450
Pro Gly Asp Leu Val Val Arg Glu Gly Ser Val Gly Arg Lys Met
455                 460                 465
Tyr Phe Ile Gln His Gly Leu Leu Ser Val Leu Ala Arg Gly Ala
470                 475                 480
Arg Asp Thr Arg Leu Thr Asp Gly Ser Tyr Phe Gly Glu Ile Cys
485                 490                 495
Leu Leu Thr Arg Gly Arg Arg Thr Ala Ser Val Arg Ala Asp Thr
500                 505                 510
Tyr Cys Arg Leu Tyr Ser Leu Ser Val Asp His Phe Asn Ala Val
```

515 520 525
Leu Glu Glu Phe Pro Met Met Arg Arg Ala Phe Glu Thr Val Ala
530 535 540
Met Asp Arg Leu Leu Arg Ile Gly Lys Lys Asn Ser Ile Leu Gln
545 550 555
Arg Lys Arg Ser Glu Pro Ser Pro Gly Ser Ser Gly Gly Ile Met
560 565 570
Glu Gln His Leu Val Gln His Asp Arg Asp Met Ala Arg Gly Val
575 580 585
Arg Gly Arg Ala Pro Ser Thr Gly Ala Gln Leu Ser Gly Lys Pro
590 595 600
Val Leu Trp Glu Pro Leu Val His Ala Pro Leu Gln Ala Ala Ala
605 610 615
Val Thr Ser Asn Val Ala Ile Ala Leu Thr His Gln Arg Gly Pro
620 625 630
Leu Pro Leu Ser Pro Asp Ser Pro Ala Thr Leu Leu Ala Arg Ser
635 640 645
Ala Trp Arg Ser Ala Gly Ser Pro Ala Ser Pro Leu Val Pro Val
650 655 660
Arg Ala Gly Pro Trp Ala Ser Thr Ser Arg Leu Pro Ala Pro Pro
665 670 675
Ala Arg Thr Leu His Ala Ser Leu Ser Arg Ala Gly Arg Ser Gln
680 685 690
Val Ser Leu Leu Gly Pro Pro Pro Gly Gly Gly Gly Arg Arg Leu
695 700 705
Gly Pro Arg Gly Arg Pro Leu Ser Ala Ser Gln Pro Ser Leu Pro
710 715 720
Gln Arg Ala Thr Gly Asp Gly Ser Pro Gly Arg Lys Gly Ser Gly
725 730 735
Ser Glu Arg Leu Pro Pro Ser Gly Leu Leu Ala Lys Pro Pro Arg
740 745 750
Thr Ala Gln Pro Pro Arg Pro Pro Val Pro Glu Pro Ala Thr Pro
755 760 765
Arg Gly Leu Gln Leu Ser Ala Asn Met
770

<210> 26
<211> 614
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5047435CD1

<400> 26
Met Ala Glu Gly Glu Arg Gly Ala Asp Val Pro His Gly Leu Gly
1 5 10 15
Ala Trp Leu Ala Asp Val Ala Leu Ala Ala Leu Arg Ala Gly Gly
20 25 30
Gln Gly Arg Arg Asp Arg Gly Gly Gly Gly Pro Glu Ser Leu Ser
35 40 45
Gly Gly Ser Gly Val Gly Asp Ser Gly Gly Gly Cys Ala Pro Gly
50 55 60
Pro Ser Ala Pro Pro Ala Arg Arg Arg Val Pro Leu Ala Met Gly
65 70 75
His Ser Pro Pro Val Leu Pro Leu Cys Ala Ser Val Ser Leu Leu
80 85 90
Gly Gly Leu Thr Phe Gly Tyr Glu Leu Ala Val Ile Ser Gly Ala
95 100 105
Leu Leu Pro Leu Gln Leu Asp Phe Gly Leu Ser Cys Leu Glu Gln
110 115 120
Glu Phe Leu Val Gly Ser Leu Leu Leu Gly Ala Leu Leu Ala Ser
125 130 135
Leu Val Gly Gly Phe Leu Ile Asp Cys Tyr Gly Arg Lys Gln Ala
140 145 150
Ile Leu Gly Ser Asn Leu Val Leu Leu Ala Gly Ser Leu Thr Leu

```
                155                 160                 165
Gly Leu Ala Gly Ser Leu Ala Trp Leu Val Leu Gly Arg Ala Val
            170                 175                 180
Val Gly Phe Ala Ile Ser Leu Ser Ser Met Ala Cys Cys Ile Tyr
        185                 190                 195
Val Ser Glu Leu Val Gly Pro Arg Gln Arg Gly Val Leu Val Ser
    200                 205                 210
Leu Tyr Glu Ala Gly Ile Thr Val Gly Ile Leu Leu Ser Tyr Ala
215                 220                 225
Leu Asn Tyr Ala Leu Ala Gly Thr Pro Trp Gly Trp Arg His Met
            230                 235                 240
Phe Gly Trp Ala Thr Ala Pro Ala Val Leu Gln Ser Leu Ser Leu
        245                 250                 255
Leu Phe Leu Pro Ala Gly Thr Asp Glu Thr Ala Thr His Lys Asp
    260                 265                 270
Leu Ile Pro Leu Gln Gly Gly Glu Ala Pro Lys Leu Gly Pro Gly
275                 280                 285
Arg Pro Arg Tyr Ser Phe Leu Asp Leu Phe Arg Ala Arg Asp Asn
            290                 295                 300
Met Arg Gly Arg Thr Thr Val Gly Leu Gly Leu Val Leu Phe Gln
        305                 310                 315
Gln Leu Thr Gly Gln Pro Asn Val Leu Cys Tyr Ala Ser Thr Ile
    320                 325                 330
Phe Ser Ser Val Gly Phe His Gly Gly Ser Ser Ala Val Leu Ala
335                 340                 345
Ser Val Gly Leu Gly Ala Val Lys Val Ala Ala Thr Leu Thr Ala
            350                 355                 360
Met Gly Leu Val Asp Arg Ala Gly Arg Arg Ala Leu Leu Leu Ala
        365                 370                 375
Gly Cys Ala Leu Met Ala Leu Ser Val Ser Gly Ile Gly Leu Val
    380                 385                 390
Ser Phe Ala Val Pro Met Asp Ser Gly Pro Ser Cys Leu Ala Val
395                 400                 405
Pro Asn Ala Thr Gly Gln Thr Gly Leu Pro Gly Asp Ser Gly Leu
            410                 415                 420
Leu Gln Asp Ser Ser Leu Pro Pro Ile Pro Arg Thr Asn Glu Asp
        425                 430                 435
Gln Arg Glu Pro Ile Leu Ser Thr Ala Lys Lys Thr Lys Pro His
    440                 445                 450
Pro Arg Ser Gly Asp Pro Ser Ala Pro Pro Arg Leu Ala Leu Ser
455                 460                 465
Ser Ala Leu Pro Gly Pro Pro Leu Pro Ala Arg Gly His Ala Leu
            470                 475                 480
Leu Arg Trp Thr Ala Leu Leu Cys Leu Met Val Phe Val Ser Ala
        485                 490                 495
Phe Ser Phe Gly Phe Gly Pro Val Thr Trp Leu Val Leu Ser Glu
    500                 505                 510
Ile Tyr Pro Val Glu Ile Arg Gly Arg Ala Phe Ala Phe Cys Asn
515                 520                 525
Ser Phe Asn Trp Ala Ala Asn Leu Phe Ile Ser Leu Ser Phe Leu
            530                 535                 540
Asp Leu Ile Gly Thr Ile Gly Leu Ser Trp Thr Phe Leu Leu Tyr
        545                 550                 555
Gly Leu Thr Ala Val Leu Gly Leu Gly Phe Ile Tyr Leu Phe Val
    560                 565                 570
Pro Glu Thr Lys Gly Gln Ser Leu Ala Glu Ile Asp Gln Gln Phe
575                 580                 585
Gln Lys Arg Arg Phe Thr Leu Ser Phe Gly His Arg Gln Asn Ser
            590                 595                 600
Thr Gly Ile Pro Tyr Ser Arg Ile Glu Ile Ser Ala Ala Ser
        605                 610
```

<210> 27
<211> 2180
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7475603CD1

<400> 27

```
Met Arg Phe Arg Lys Gly Gln Glu Leu Pro Ala Ala Ala Pro His
  1               5                  10                  15
Val Phe Ser Pro Thr Val Val Leu Thr Ser Leu Ser Arg Pro Leu
             20                  25                  30
Pro Ser Leu Thr Met Ala Phe Trp Thr Gln Leu Met Leu Leu Leu
         35                  40                  45
Trp Lys Asn Phe Met Tyr Arg Arg Arg Gln Pro Val Gln Leu Leu
     50                  55                  60
Val Glu Leu Leu Trp Pro Leu Phe Leu Phe Phe Ile Leu Val Ala
 65                  70                  75
Val Arg His Ser His Pro Pro Leu Glu His His Glu Cys His Phe
             80                  85                  90
Pro Asn Lys Pro Leu Pro Ser Ala Gly Thr Val Pro Trp Leu Gln
         95                 100                 105
Gly Leu Ile Cys Asn Val Asn Asn Thr Cys Phe Pro Gln Leu Thr
    110                 115                 120
Pro Gly Glu Glu Pro Gly Arg Leu Ser Asn Phe Asn Asp Ser Leu
125                 130                 135
Val Ser Arg Leu Leu Ala Asp Ala Arg Thr Val Leu Gly Gly Ala
            140                 145                 150
Ser Ala His Arg Thr Leu Ala Gly Leu Gly Lys Leu Ile Ala Thr
        155                 160                 165
Leu Arg Ala Ala Arg Ser Thr Ala Gln Pro Gln Pro Thr Lys Gln
    170                 175                 180
Ser Pro Leu Glu Pro Pro Met Leu Asp Val Ala Glu Leu Leu Thr
185                 190                 195
Ser Leu Leu Arg Thr Glu Ser Leu Gly Leu Ala Leu Gly Gln Ala
            200                 205                 210
Gln Glu Pro Leu His Ser Leu Leu Glu Ala Ala Glu Asp Leu Ala
        215                 220                 225
Gln Glu Leu Leu Ala Leu Arg Ser Leu Val Glu Leu Arg Ala Leu
    230                 235                 240
Leu Gln Arg Pro Arg Gly Thr Ser Gly Pro Leu Glu Leu Leu Ser
245                 250                 255
Glu Ala Leu Cys Ser Val Arg Gly Pro Ser Ser Thr Val Gly Pro
            260                 265                 270
Ser Leu Asn Trp Tyr Glu Ala Ser Asp Leu Met Glu Leu Val Gly
        275                 280                 285
Gln Glu Pro Glu Ser Ala Leu Pro Asp Ser Ser Leu Ser Pro Ala
    290                 295                 300
Cys Ser Glu Leu Ile Gly Ala Leu Asp Ser His Pro Leu Ser Arg
305                 310                 315
Leu Leu Trp Arg Arg Leu Lys Pro Leu Ile Leu Gly Lys Leu Leu
            320                 325                 330
Phe Ala Pro Asp Thr Pro Phe Thr Arg Lys Leu Met Ala Gln Val
        335                 340                 345
Asn Arg Thr Phe Glu Glu Leu Thr Leu Leu Arg Asp Val Arg Glu
    350                 355                 360
Val Trp Glu Met Leu Gly Pro Arg Ile Phe Thr Phe Met Asn Asp
365                 370                 375
Ser Ser Asn Val Ala Met Leu Gln Arg Leu Leu Gln Met Gln Asp
            380                 385                 390
Glu Gly Arg Arg Gln Pro Arg Pro Gly Gly Arg Asp His Met Glu
        395                 400                 405
Ala Leu Arg Ser Phe Leu Asp Pro Gly Ser Gly Gly Tyr Ser Trp
    410                 415                 420
Gln Asp Ala His Ala Asp Val Gly His Leu Val Gly Thr Leu Gly
425                 430                 435
Arg Val Thr Glu Cys Leu Ser Leu Asp Lys Leu Glu Ala Ala Pro
            440                 445                 450
Ser Glu Ala Ala Leu Val Ser Arg Ala Leu Gln Leu Leu Ala Glu
        455                 460                 465
```

```
His Arg Phe Trp Ala Gly Val Val Phe Leu Gly Pro Glu Asp Ser
470                 475                 480
Ser Asp Pro Thr Glu His Pro Thr Pro Asp Leu Gly Pro Gly His
485                 490                 495
Val Arg Ile Lys Ile Arg Met Asp Ile Asp Val Val Thr Arg Thr
500                 505                 510
Asn Lys Ile Arg Asp Arg Phe Trp Asp Pro Gly Pro Ala Ala Asp
515                 520                 525
Pro Leu Thr Asp Leu Arg Tyr Val Trp Gly Gly Phe Val Tyr Leu
530                 535                 540
Gln Asp Leu Val Glu Arg Ala Ala Val Arg Val Leu Ser Gly Ala
545                 550                 555
Asn Pro Arg Ala Gly Leu Tyr Leu Gln Gln Met Pro Tyr Pro Cys
560                 565                 570
Tyr Val Asp Asp Val Phe Leu Arg Val Leu Ser Arg Ser Leu Pro
575                 580                 585
Leu Phe Leu Thr Leu Ala Trp Ile Tyr Ser Val Thr Leu Thr Val
590                 595                 600
Lys Ala Val Val Arg Glu Lys Glu Thr Arg Leu Arg Asp Thr Met
605                 610                 615
Arg Ala Met Gly Leu Ser Arg Ala Val Leu Trp Leu Gly Trp Phe
620                 625                 630
Leu Ser Cys Leu Gly Pro Phe Leu Leu Ser Ala Ala Leu Leu Val
635                 640                 645
Leu Val Leu Lys Leu Gly Asp Ile Leu Pro Tyr Ser His Pro Gly
650                 655                 660
Val Val Phe Leu Phe Leu Ala Ala Phe Ala Val Ala Thr Val Thr
665                 670                 675
Gln Ser Phe Leu Leu Ser Ala Phe Phe Ser Arg Ala Asn Leu Ala
680                 685                 690
Ala Ala Cys Gly Gly Leu Ala Tyr Phe Ser Leu Tyr Leu Pro Tyr
695                 700                 705
Val Leu Cys Val Ala Trp Arg Asp Arg Leu Pro Ala Gly Gly Arg
710                 715                 720
Val Ala Ala Ser Leu Leu Ser Pro Val Ala Phe Gly Phe Gly Cys
725                 730                 735
Glu Ser Leu Ala Leu Leu Glu Glu Gln Gly Glu Gly Ala Gln Trp
740                 745                 750
His Asn Val Gly Thr Arg Pro Thr Ala Asp Val Phe Ser Leu Ala
755                 760                 765
Gln Val Ser Gly Leu Leu Leu Leu Asp Ala Ala Leu Tyr Gly Leu
770                 775                 780
Ala Thr Trp Tyr Leu Glu Ala Val Cys Pro Gly Gln Tyr Gly Ile
785                 790                 795
Pro Glu Pro Trp Asn Phe Pro Phe Arg Arg Ser Tyr Trp Cys Gly
800                 805                 810
Pro Arg Pro Pro Lys Ser Pro Ala Pro Cys Pro Thr Pro Leu Asp
815                 820                 825
Pro Lys Val Leu Val Glu Glu Ala Pro Pro Gly Leu Ser Pro Gly
830                 835                 840
Val Ser Val Arg Ser Leu Glu Lys Arg Phe Pro Gly Ser Pro Gln
845                 850                 855
Pro Ala Leu Arg Gly Leu Ser Leu Asp Phe Tyr Gln Gly His Ile
860                 865                 870
Thr Ala Phe Leu Gly His Asn Gly Ala Gly Lys Thr Thr Thr Leu
875                 880                 885
Ser Ile Leu Ser Gly Leu Phe Pro Pro Ser Gly Gly Ser Ala Phe
890                 895                 900
Ile Leu Gly His Asp Val Arg Ser Ser Met Ala Ala Ile Arg Pro
905                 910                 915
His Leu Gly Val Cys Pro Gln Tyr Asn Val Leu Phe Asp Met Leu
920                 925                 930
Thr Val Asp Glu His Val Trp Phe Tyr Gly Arg Leu Lys Gly Leu
935                 940                 945
Ser Ala Ala Val Val Gly Pro Glu Gln Asp Arg Leu Leu Gln Asp
950                 955                 960
Val Gly Leu Val Ser Lys Gln Ser Val Gln Thr Arg His Leu Ser
```

965 970 975
Gly Gly Met Gln Arg Lys Leu Ser Val Ala Ile Ala Phe Val Gly
980 985 990
Gly Ser Gln Val Val Ile Leu Asp Glu Pro Thr Ala Gly Val Asp
995 1000 1005
Pro Ala Ser Arg Arg Gly Ile Trp Glu Leu Leu Leu Lys Tyr Arg
1010 1015 1020
Glu Gly Arg Thr Leu Ile Leu Ser Thr His His Leu Asp Glu Ala
1025 1030 1035
Glu Leu Leu Gly Asp Arg Val Ala Val Val Ala Gly Gly Arg Leu
1040 1045 1050
Cys Cys Cys Gly Ser Pro Leu Phe Leu Arg Arg His Leu Gly Ser
1055 1060 1065
Gly Tyr Tyr Leu Thr Leu Val Lys Ala Arg Leu Pro Leu Thr Thr
1070 1075 1080
Asn Glu Lys Ala Asp Thr Asp Met Glu Gly Ser Val Asp Thr Arg
1085 1090 1095
Gln Glu Lys Lys Asn Gly Ser Gln Gly Ser Arg Val Gly Thr Pro
1100 1105 1110
Gln Leu Leu Ala Leu Val Gln His Trp Val Pro Gly Ala Arg Leu
1115 1120 1125
Val Glu Glu Leu Pro His Glu Leu Val Leu Val Leu Pro Tyr Thr
1130 1135 1140
Gly Ala His Asp Gly Ser Phe Ala Thr Leu Phe Arg Glu Leu Asp
1145 1150 1155
Thr Arg Leu Ala Glu Leu Arg Leu Thr Gly Tyr Gly Ile Ser Asp
1160 1165 1170
Thr Ser Leu Glu Glu Ile Phe Leu Lys Val Val Glu Glu Cys Ala
1175 1180 1185
Ala Asp Thr Asp Met Glu Asp Gly Ser Cys Gly Gln His Leu Cys
1190 1195 1200
Thr Gly Ile Ala Gly Leu Asp Val Thr Leu Arg Leu Lys Met Pro
1205 1210 1215
Pro Gln Glu Thr Ala Leu Glu Asn Gly Glu Pro Ala Gly Ser Ala
1220 1225 1230
Pro Glu Thr Asp Gln Gly Ser Gly Pro Asp Ala Val Gly Arg Val
1235 1240 1245
Gln Gly Trp Ala Leu Thr Arg Gln Gln Leu Gln Ala Leu Leu Leu
1250 1255 1260
Lys Arg Phe Leu Leu Ala Arg Arg Ser Arg Arg Gly Leu Phe Ala
1265 1270 1275
Gln Ile Val Leu Pro Ala Leu Phe Val Gly Leu Ala Leu Val Phe
1280 1285 1290
Ser Leu Ile Val Pro Pro Phe Gly His Tyr Pro Ala Leu Arg Leu
1295 1300 1305
Ser Pro Thr Met Tyr Gly Ala Gln Val Ser Phe Phe Ser Glu Asp
1310 1315 1320
Ala Pro Gly Asp Pro Gly Arg Ala Arg Leu Leu Glu Ala Leu Leu
1325 1330 1335
Gln Glu Ala Gly Leu Glu Glu Pro Pro Val Gln His Ser Ser His
1340 1345 1350
Arg Phe Ser Ala Pro Glu Val Pro Ala Glu Val Ala Lys Val Leu
1355 1360 1365
Ala Ser Gly Asn Trp Thr Pro Glu Ser Pro Ser Pro Ala Cys Gln
1370 1375 1380
Cys Ser Arg Pro Gly Ala Arg Arg Leu Leu Pro Asp Cys Pro Ala
1385 1390 1395
Ala Ala Gly Gly Pro Pro Pro Pro Gln Ala Val Thr Gly Ser Gly
1400 1405 1410
Glu Val Val Gln Asn Gln Thr Gly Arg Asn Leu Ser Asp Phe Leu
1415 1420 1425
Val Lys Thr Tyr Pro Arg Leu Val Arg Gln Gly Leu Lys Thr Lys
1430 1435 1440
Lys Trp Val Asn Glu Val Arg Tyr Gly Gly Phe Ser Leu Gly Gly
1445 1450 1455
Arg Asp Pro Gly Leu Pro Ser Gly Gln Glu Leu Gly Arg Ser Val
1460 1465 1470

Glu Glu Leu Trp Ala Leu Leu Ser Pro Leu Pro Gly Gly Ala Leu
1475 1480 1485
Asp Arg Val Leu Lys Asn Leu Thr Ala Trp Ala His Ser Leu Asp
1490 1495 1500
Ala Gln Asp Ser Leu Lys Ile Trp Phe Asn Asn Lys Gly Trp His
1505 1510 1515
Ser Met Val Ala Phe Val Asn Arg Ala Ser Asn Ala Ile Leu Arg
1520 1525 1530
Ala His Leu Pro Pro Gly Pro Ala Arg His Ala His Ser Ile Thr
1535 1540 1545
Thr Leu Asn His Pro Leu Asn Leu Thr Lys Glu Gln Leu Ser Glu
1550 1555 1560
Ala Ala Leu Met Ala Ser Ser Val Asp Val Leu Val Ser Ile Cys
1565 1570 1575
Val Val Phe Ala Met Ser Phe Val Pro Ala Ser Phe Thr Leu Val
1580 1585 1590
Leu Ile Glu Glu Arg Val Thr Arg Ala Lys His Leu Gln Leu Met
1595 1600 1605
Gly Gly Leu Ser Pro Thr Leu Tyr Trp Leu Gly Asn Phe Leu Trp
1610 1615 1620
Asp Met Cys Asn Tyr Leu Val Pro Ala Cys Ile Val Val Leu Ile
1625 1630 1635
Phe Leu Ala Phe Gln Gln Arg Ala Tyr Val Ala Pro Ala Asn Leu
1640 1645 1650
Pro Ala Leu Leu Leu Leu Leu Leu Leu Tyr Gly Trp Ser Ile Thr
1655 1660 1665
Pro Leu Met Tyr Pro Ala Ser Phe Phe Phe Ser Val Pro Ser Thr
1670 1675 1680
Ala Tyr Val Val Leu Thr Cys Ile Asn Leu Phe Ile Gly Ile Asn
1685 1690 1695
Gly Ser Met Ala Thr Phe Val Leu Glu Leu Phe Ser Asp Gln Lys
1700 1705 1710
Leu Gln Glu Val Ser Arg Ile Leu Lys Gln Val Phe Leu Ile Phe
1715 1720 1725
Pro His Phe Cys Leu Gly Arg Gly Leu Ile Asp Met Val Arg Asn
1730 1735 1740
Gln Ala Met Ala Asp Ala Phe Glu Arg Leu Gly Asp Arg Gln Phe
1745 1750 1755
Gln Ser Pro Leu Arg Trp Glu Val Val Gly Lys Asn Leu Leu Ala
1760 1765 1770
Met Val Ile Gln Gly Pro Leu Phe Leu Leu Phe Thr Leu Leu Leu
1775 1780 1785
Gln His Arg Ser Gln Leu Leu Pro Gln Pro Arg Val Arg Ser Leu
1790 1795 1800
Pro Leu Leu Gly Glu Glu Asp Glu Asp Val Ala Arg Glu Arg Glu
1805 1810 1815
Arg Val Val Gln Gly Ala Thr Gln Gly Asp Val Leu Val Leu Arg
1820 1825 1830
Asn Leu Thr Lys Val Tyr Arg Gly Gln Arg Met Pro Ala Val Asp
1835 1840 1845
Arg Leu Cys Leu Gly Ile Pro Pro Gly Glu Cys Phe Gly Leu Leu
1850 1855 1860
Gly Val Asn Gly Ala Gly Lys Thr Ser Thr Phe Arg Met Val Thr
1865 1870 1875
Gly Asp Thr Leu Ala Ser Arg Gly Glu Ala Val Leu Ala Gly His
1880 1885 1890
Ser Val Ala Arg Glu Pro Ser Ala Ala His Leu Ser Met Gly Tyr
1895 1900 1905
Cys Pro Gln Ser Asp Ala Ile Phe Glu Leu Leu Thr Gly Arg Glu
1910 1915 1920
His Leu Glu Leu Leu Ala Arg Leu Arg Gly Val Pro Glu Ala Gln
1925 1930 1935
Val Ala Gln Thr Ala Gly Ser Gly Leu Ala Arg Leu Gly Leu Ser
1940 1945 1950
Trp Tyr Ala Asp Arg Pro Ala Gly Thr Tyr Ser Gly Gly Asn Lys
1955 1960 1965
Arg Lys Leu Ala Thr Ala Leu Ala Leu Val Gly Asp Pro Ala Val

```
                1970                1975                1980
Val Phe Leu Asp Glu Pro Thr Thr Gly Met Asp Pro Ser Ala Arg
                1985                1990                1995
Arg Phe Leu Trp Asn Ser Leu Leu Ala Val Val Arg Glu Gly Arg
                2000                2005                2010
Ser Val Met Leu Thr Ser His Ser Met Glu Glu Cys Glu Ala Leu
                2015                2020                2025
Cys Ser Arg Leu Ala Ile Met Val Asn Gly Arg Phe Arg Cys Leu
                2030                2035                2040
Gly Ser Pro Gln His Leu Lys Gly Arg Phe Ala Ala Gly His Thr
                2045                2050                2055
Leu Thr Leu Arg Val Pro Ala Ala Arg Ser Gln Pro Ala Ala Ala
                2060                2065                2070
Phe Val Ala Ala Glu Phe Pro Gly Ala Glu Leu Arg Glu Ala His
                2075                2080                2085
Gly Gly Arg Leu Arg Phe Gln Leu Pro Pro Gly Gly Arg Cys Ala
                2090                2095                2100
Leu Ala Arg Val Phe Gly Glu Leu Ala Val His Gly Ala Glu His
                2105                2110                2115
Gly Val Glu Asp Phe Ser Val Ser Gln Thr Met Leu Glu Glu Val
                2120                2125                2130
Phe Leu Tyr Phe Ser Lys Asp Gln Gly Lys Asp Glu Asp Thr Glu
                2135                2140                2145
Glu Gln Lys Glu Ala Gly Val Gly Val Asp Pro Ala Pro Gly Leu
                2150                2155                2160
Gln His Pro Lys Arg Val Ser Gln Phe Leu Asp Asp Pro Ser Thr
                2165                2170                2175
Ala Glu Thr Val Leu
                2180
```

<210> 28
<211> 1737
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477845CD1

<400> 28

```
Met Leu Lys Arg Lys Gln Ser Ser Arg Val Glu Ala Gln Pro Val
  1                 5                  10                  15
Thr Asp Phe Gly Pro Asp Glu Ser Leu Ser Asp Asn Ala Asp Ile
                 20                  25                  30
Leu Trp Ile Asn Lys Pro Trp Val His Ser Leu Leu Arg Ile Cys
                 35                  40                  45
Ala Ile Ile Ser Val Ile Ser Val Cys Met Asn Thr Pro Met Thr
                 50                  55                  60
Phe Glu His Tyr Pro Pro Leu Gln Tyr Val Thr Phe Thr Leu Asp
                 65                  70                  75
Thr Leu Leu Met Phe Leu Tyr Thr Ala Glu Met Ile Ala Lys Met
                 80                  85                  90
His Ile Arg Gly Ile Val Lys Gly Asp Ser Ser Tyr Val Lys Asp
                 95                 100                 105
Arg Trp Cys Val Phe Asp Gly Phe Met Val Phe Cys Leu Trp Val
                110                 115                 120
Ser Leu Val Leu Gln Val Phe Glu Ile Ala Asp Ile Val Asp Gln
                125                 130                 135
Met Ser Pro Trp Gly Met Leu Arg Ile Pro Arg Pro Leu Ile Met
                140                 145                 150
Ile Arg Ala Phe Arg Ile Tyr Phe Arg Phe Glu Leu Pro Arg Thr
                155                 160                 165
Arg Ile Thr Asn Ile Leu Lys Arg Ser Gly Glu Gln Ile Trp Ser
                170                 175                 180
Val Ser Ile Phe Leu Leu Phe Phe Leu Leu Leu Tyr Gly Ile Leu
                185                 190                 195
Gly Val Gln Met Phe Gly Thr Phe Thr Tyr His Cys Val Val Asn
```

```
                200                 205                 210
Asp Thr Lys Pro Gly Asn Val Thr Trp Asn Ser Leu Ala Ile Pro
                215                 220                 225
Asp Thr His Cys Ser Pro Glu Leu Glu Glu Gly Tyr Gln Cys Pro
                230                 235                 240
Pro Gly Phe Lys Cys Met Asp Leu Glu Asp Leu Gly Leu Ser Arg
                245                 250                 255
Gln Glu Leu Gly Tyr Ser Gly Phe Asn Glu Ile Gly Thr Ser Ile
                260                 265                 270
Phe Thr Val Tyr Glu Ala Ala Ser Gln Glu Gly Trp Val Phe Leu
                275                 280                 285
Met Tyr Arg Ala Ile Asp Ser Phe Pro Arg Trp Arg Ser Tyr Phe
                290                 295                 300
Tyr Phe Ile Thr Leu Ile Phe Phe Leu Ala Trp Leu Val Lys Asn
                305                 310                 315
Val Phe Ile Ala Val Ile Ile Glu Thr Phe Ala Glu Ile Arg Val
                320                 325                 330
Gln Phe Gln Gln Met Trp Gly Ser Arg Ser Ser Thr Thr Ser Thr
                335                 340                 345
Ala Thr Thr Gln Met Phe His Glu Asp Ala Ala Gly Gly Trp Gln
                350                 355                 360
Leu Val Ala Val Asp Val Asn Lys Pro Gln Gly Arg Ala Pro Ala
                365                 370                 375
Cys Leu Gln Lys Met Met Arg Ser Ser Val Phe His Met Phe Ile
                380                 385                 390
Leu Ser Met Val Thr Val Asp Val Ile Val Ala Ala Ser Asn Tyr
                395                 400                 405
Tyr Lys Gly Glu Asn Phe Arg Arg Gln Tyr Asp Glu Phe Tyr Leu
                410                 415                 420
Ala Glu Val Ala Phe Thr Val Leu Phe Asp Leu Glu Ala Leu Leu
                425                 430                 435
Lys Ile Trp Cys Leu Gly Phe Thr Gly Tyr Ile Ser Ser Ser Leu
                440                 445                 450
His Lys Phe Glu Leu Leu Leu Val Ile Gly Thr Thr Leu His Val
                455                 460                 465
Tyr Pro Asp Leu Tyr His Ser Gln Phe Thr Tyr Phe Gln Val Leu
                470                 475                 480
Arg Val Val Arg Leu Ile Lys Ile Ser Pro Ala Leu Glu Asp Phe
                485                 490                 495
Val Tyr Lys Ile Phe Gly Pro Gly Lys Lys Leu Gly Ser Leu Val
                500                 505                 510
Val Phe Thr Ala Ser Leu Leu Ile Val Met Ser Ala Ile Ser Leu
                515                 520                 525
Gln Met Phe Cys Phe Val Glu Glu Leu Asp Arg Phe Thr Thr Phe
                530                 535                 540
Pro Arg Ala Phe Met Ser Met Phe Gln Ile Leu Thr Gln Glu Gly
                545                 550                 555
Trp Val Asp Val Met Asp Gln Thr Leu Asn Ala Val Gly His Met
                560                 565                 570
Trp Ala Pro Val Val Ala Ile Tyr Phe Ile Leu Tyr His Leu Phe
                575                 580                 585
Ala Thr Leu Ile Leu Leu Ser Leu Phe Val Ala Val Ile Leu Asp
                590                 595                 600
Asn Leu Glu Leu Asp Glu Asp Leu Lys Lys Leu Lys Gln Leu Lys
                605                 610                 615
Gln Ser Glu Ala Asn Ala Asp Thr Lys Glu Lys Leu Pro Leu Arg
                620                 625                 630
Leu Arg Ile Phe Glu Lys Phe Pro Asn Arg Pro Gln Met Val Lys
                635                 640                 645
Ile Ser Lys Leu Pro Ser Asp Phe Thr Val Pro Lys Ile Arg Glu
                650                 655                 660
Ser Phe Met Lys Gln Phe Ile Asp Arg Gln Gln Gln Asp Thr Cys
                665                 670                 675
Cys Leu Leu Arg Ser Leu Pro Thr Thr Ser Ser Ser Ser Cys Asp
                680                 685                 690
His Ser Lys Arg Ser Ala Ile Glu Asp Asn Lys Tyr Ile Asp Gln
                695                 700                 705
```

Lys Leu Arg Lys Ser Val Phe Ser Ile Arg Ala Arg Asn Leu Leu
710 715 720
Glu Lys Glu Thr Ala Val Thr Lys Ile Leu Arg Ala Cys Thr Arg
725 730 735
Gln Arg Met Leu Ser Gly Ser Phe Glu Gly Gln Pro Ala Lys Glu
740 745 750
Arg Ser Ile Leu Ser Val Gln His His Ile Arg Gln Glu Arg Arg
755 760 765
Ser Leu Arg His Gly Ser Asn Ser Gln Arg Ile Ser Arg Gly Lys
770 775 780
Ser Leu Glu Thr Leu Thr Gln Asp His Cys Asn Thr Val Ile Tyr
785 790 795
Arg Asn Ala Gln Arg Glu Val Ser Glu Ile Lys Met Ile Gln Glu
800 805 810
Lys Lys Glu Leu Ala Glu Met Leu Gln Gly Lys Cys Lys Lys Glu
815 820 825
Leu Arg Glu Ser His Pro Tyr Phe Asp Lys Pro Leu Phe Ile Val
830 835 840
Gly Arg Glu His Arg Phe Arg Asn Phe Cys Arg Val Val Val Arg
845 850 855
Ala Arg Phe Asn Ala Ser Lys Thr Asp Pro Val Thr Gly Ala Val
860 865 870
Lys Asn Thr Lys Tyr His Leu Leu Tyr Asp Leu Leu Gly Leu Val
875 880 885
Thr Tyr Leu Asp Trp Val Met Ile Ile Val Thr Ser Asp Ser Cys
890 895 900
Ile Ser Met Met Phe Glu Ser Pro Phe Arg Arg Val Met His Ala
905 910 915
Pro Thr Leu Gln Ile Ala Glu Tyr Val Phe Val Ile Phe Met Ser
920 925 930
Ile Glu Leu Asn Leu Lys Ile Met Ala Asp Gly Leu Phe Phe Thr
935 940 945
Pro Thr Ala Val Ile Arg Asp Phe Gly Gly Val Met Asp Ile Phe
950 955 960
Ile Tyr Leu Val Ser Leu Ile Phe Leu Cys Trp Met Pro Gln Asn
965 970 975
Val Pro Ala Glu Ser Gly Ala Gln Leu Leu Met Val Leu Arg Cys
980 985 990
Leu Arg Pro Leu Arg Ile Phe Lys Leu Val Pro Gln Met Arg Lys
995 1000 1005
Val Val Arg Glu Leu Phe Ser Gly Phe Lys Glu Ile Phe Leu Val
1010 1015 1020
Ser Ile Leu Leu Leu Thr Leu Met Leu Val Phe Ala Ser Phe Gly
1025 1030 1035
Val Gln Leu Phe Ala Gly Lys Leu Ala Lys Cys Asn Asp Pro Asn
1040 1045 1050
Ile Ile Arg Arg Glu Asp Cys Asn Gly Ile Phe Arg Ile Asn Val
1055 1060 1065
Ser Val Ser Lys Asn Leu Asn Leu Lys Leu Arg Pro Gly Glu Lys
1070 1075 1080
Lys Pro Gly Phe Trp Val Pro Arg Val Trp Ala Asn Pro Arg Asn
1085 1090 1095
Phe Asn Phe Asp Asn Val Gly Asn Ala Met Leu Ala Leu Phe Glu
1100 1105 1110
Val Leu Ser Leu Lys Gly Trp Val Glu Val Arg Asp Val Ile Ile
1115 1120 1125
His Arg Val Gly Pro Ile His Gly Ile Tyr Ile His Val Phe Val
1130 1135 1140
Phe Leu Gly Cys Met Ile Gly Leu Thr Leu Phe Val Gly Val Val
1145 1150 1155
Ile Ala Asn Phe Asn Glu Asn Lys Gly Thr Ala Leu Leu Thr Val
1160 1165 1170
Asp Gln Arg Arg Trp Glu Asp Leu Lys Ser Arg Leu Lys Ile Ala
1175 1180 1185
Gln Pro Leu His Leu Pro Pro Arg Pro Asp Asn Asp Gly Phe Arg
1190 1195 1200
Ala Lys Met Tyr Asp Ile Thr Gln His Pro Phe Phe Lys Arg Thr

1205 1210 1215
Ile Ala Leu Leu Val Leu Ala Gln Ser Val Leu Leu Ser Val Lys
1220 1225 1230
Trp Asp Val Glu Asp Pro Val Thr Val Pro Leu Ala Thr Met Ser
1235 1240 1245
Val Val Phe Thr Phe Ile Phe Val Leu Glu Val Thr Met Lys Ile
1250 1255 1260
Ile Ala Met Ser Pro Ala Gly Phe Trp Gln Ser Arg Arg Asn Arg
1265 1270 1275
Tyr Asp Leu Leu Val Thr Ser Leu Gly Val Val Trp Val Val Leu
1280 1285 1290
His Phe Ala Leu Leu Asn Ala Tyr Thr Tyr Met Met Gly Ala Cys
1295 1300 1305
Val Ile Val Phe Arg Phe Phe Ser Ile Cys Gly Lys His Val Thr
1310 1315 1320
Leu Lys Met Leu Leu Leu Thr Val Val Val Ser Met Tyr Lys Ser
1325 1330 1335
Phe Phe Ile Ile Val Gly Met Phe Leu Leu Leu Leu Cys Tyr Ala
1340 1345 1350
Phe Ala Gly Val Val Leu Phe Gly Thr Val Lys Tyr Gly Glu Asn
1355 1360 1365
Ile Asn Arg His Ala Asn Phe Ser Ser Ala Gly Lys Ala Ile Thr
1370 1375 1380
Val Leu Phe Arg Ile Val Thr Gly Glu Asp Trp Asn Lys Ile Met
1385 1390 1395
His Asp Cys Met Val Gln Pro Pro Phe Cys Thr Pro Asp Glu Phe
1400 1405 1410
Thr Tyr Trp Ala Thr Asp Cys Gly Asn Tyr Ala Gly Ala Leu Met
1415 1420 1425
Tyr Phe Cys Ser Phe Tyr Val Ile Ile Ala Tyr Ile Met Leu Asn
1430 1435 1440
Leu Leu Val Ala Ile Ile Val Glu Asn Phe Ser Leu Ile Tyr Ser
1445 1450 1455
Thr Glu Glu Asp Gln Leu Leu Ser Tyr Asn Asp Leu Arg His Phe
1460 1465 1470
Gln Ile Ile Trp Asn Met Val Asp Asp Lys Arg Glu Val Phe Pro
1475 1480 1485
Thr Phe Arg Val Lys Phe Leu Leu Arg Leu Leu Arg Gly Arg Leu
1490 1495 1500
Glu Val Asp Leu Asp Lys Asp Lys Leu Leu Phe Lys His Met Cys
1505 1510 1515
Tyr Glu Met Glu Arg Leu His Asn Gly Gly Asp Val Thr Phe His
1520 1525 1530
Asp Val Leu Ser Met Leu Ser Tyr Arg Ser Val Asp Ile Arg Lys
1535 1540 1545
Ser Leu Gln Leu Glu Glu Leu Leu Ala Arg Glu Gln Leu Glu Tyr
1550 1555 1560
Thr Ile Glu Glu Glu Val Ala Lys Gln Thr Ile Arg Met Trp Leu
1565 1570 1575
Lys Lys Cys Leu Lys Arg Ile Arg Ala Lys Gln Gln Gln Ser Cys
1580 1585 1590
Ser Ile Ile His Ser Leu Arg Glu Ser Gln Gln Gln Glu Leu Ser
1595 1600 1605
Arg Phe Leu Asn Pro Pro Ser Ile Glu Thr Thr Gln Pro Ser Glu
1610 1615 1620
Asp Thr Asn Ala Asn Ser Gln Asp Asn Ser Met Gln Pro Glu Thr
1625 1630 1635
Ser Ser Gln Gln Gln Leu Leu Ser Pro Thr Leu Ser Asp Arg Gly
1640 1645 1650
Gly Ser Arg Gln Asp Ala Ala Asp Ala Gly Lys Pro Gln Arg Lys
1655 1660 1665
Phe Gly Gln Trp Arg Leu Pro Ser Ala Pro Lys Pro Ile Ser His
1670 1675 1680
Ser Val Ser Ser Val Asn Leu Arg Phe Gly Gly Arg Thr Thr Met
1685 1690 1695
Lys Ser Val Val Cys Lys Met Asn Pro Met Thr Asp Ala Ala Ser
1700 1705 1710

Cys Gly Ser Glu Val Lys Lys Trp Trp Thr Arg Gln Leu Thr Val
1715 1720 1725
Glu Ser Asp Glu Ser Gly Asp Asp Leu Leu Asp Ile
1730 1735

<210> 29
<211> 547
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 168827CD1

<400> 29
Met Ala Phe Gln Asp Leu Leu Asp Gln Val Gly Gly Leu Gly Arg
1 5 10 15
Phe Gln Ile Leu Gln Met Val Phe Leu Ile Met Phe Asn Val Ile
20 25 30
Val Tyr His Gln Thr Gln Leu Glu Asn Phe Ala Ala Phe Ile Leu
35 40 45
Asp His Arg Cys Trp Val His Ile Leu Asp Asn Asp Thr Ile Pro
50 55 60
Asp Asn Asp Pro Gly Thr Leu Ser Gln Asp Ala Leu Leu Arg Ile
65 70 75
Ser Ile Pro Phe Asp Ser Asn Leu Arg Pro Glu Lys Cys Arg Arg
80 85 90
Phe Val His Pro Gln Trp Lys Leu Ile His Leu Asn Gly Thr Phe
95 100 105
Pro Asn Thr Ser Glu Pro Asp Thr Glu Pro Cys Val Asp Gly Trp
110 115 120
Val Tyr Asp Gln Ser Ser Phe Pro Ser Thr Ile Val Thr Lys Trp
125 130 135
Asp Leu Val Cys Glu Ser Gln Pro Leu Asn Ser Val Ala Lys Phe
140 145 150
Leu Phe Met Ala Gly Met Met Val Gly Gly Asn Leu Tyr Gly His
155 160 165
Leu Ser Asp Arg Phe Gly Arg Lys Phe Val Leu Arg Trp Ser Tyr
170 175 180
Leu Gln Leu Ala Ile Val Gly Thr Cys Ala Ala Phe Ala Pro Thr
185 190 195
Ile Leu Val Tyr Cys Ser Leu Arg Phe Leu Ala Gly Ala Ala Thr
200 205 210
Phe Ser Ile Ile Val Asn Thr Val Leu Leu Ile Val Glu Trp Ile
215 220 225
Thr His Gln Phe Cys Ala Met Ala Leu Thr Leu Thr Leu Cys Ala
230 235 240
Ala Ser Ile Gly His Ile Thr Leu Gly Ser Leu Ala Phe Val Ile
245 250 255
Arg Asp Gln Cys Ile Leu Gln Leu Val Met Ser Ala Pro Cys Phe
260 265 270
Val Phe Phe Leu Phe Ser Arg Trp Leu Ala Glu Ser Ala Arg Trp
275 280 285
Leu Ile Ile Asn Asn Lys Pro Glu Glu Gly Leu Lys Glu Leu Thr
290 295 300
Lys Ala Ala His Arg Asn Gly Met Lys Asn Ala Glu Asp Ile Leu
305 310 315
Thr Met Glu Val Leu Lys Ser Thr Met Lys Gln Glu Leu Glu Ala
320 325 330
Ala Gln Lys Lys His Ser Leu Cys Glu Leu Leu Arg Ile Pro Asn
335 340 345
Ile Cys Lys Arg Ile Cys Phe Leu Ser Phe Val Arg Phe Ala Ser
350 355 360
Thr Ile Pro Phe Trp Gly Leu Thr Leu His Leu Gln His Leu Gly
365 370 375
Asn Asn Val Phe Leu Leu Gln Thr Leu Phe Gly Ala Val Thr Leu
380 385 390

```
Leu Ala Asn Cys Val Ala Pro Trp Ala Leu Asn His Met Ser Arg
                395                 400                 405
Arg Leu Ser Gln Met Leu Leu Met Phe Leu Leu Ala Thr Cys Leu
                410                 415                 420
Leu Ala Ile Ile Phe Val Pro Gln Glu Met Gln Thr Leu Arg Val
                425                 430                 435
Val Leu Ala Thr Leu Gly Val Gly Ala Ala Ser Leu Gly Ile Thr
                440                 445                 450
Cys Ser Thr Ala Gln Glu Asn Glu Leu Ile Pro Ser Ile Ile Arg
                455                 460                 465
Gly Arg Ala Thr Gly Ile Thr Gly Asn Phe Ala Asn Ile Gly Gly
                470                 475                 480
Ala Leu Ala Ser Leu Met Met Ile Leu Ser Ile Tyr Ser Arg Pro
                485                 490                 495
Leu Pro Trp Ile Ile Tyr Gly Val Phe Ala Ile Leu Ser Gly Leu
                500                 505                 510
Val Val Leu Leu Leu Pro Glu Thr Arg Asn Gln Pro Leu Leu Asp
                515                 520                 525
Ser Ile Gln Asp Val Glu Asn Glu Gly Val Asn Ser Leu Ala Ala
                530                 535                 540
Pro Gln Arg Ser Ser Val Leu
                545
```

<210> 30
<211> 547
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472734CD1

<400> 30

```
Met Gly Phe Asp Val Leu Leu Asp Gln Val Gly Gly Met Gly Arg
  1                 5                  10                  15
Phe Gln Ile Cys Leu Ile Ala Phe Phe Cys Ile Thr Asn Ile Leu
                 20                  25                  30
Leu Phe Pro Asn Ile Val Leu Glu Asn Phe Thr Ala Phe Thr Pro
                 35                  40                  45
Ser His Arg Cys Trp Val Pro Leu Leu Asp Asn Asp Thr Val Ser
                 50                  55                  60
Asp Asn Asp Thr Gly Thr Leu Ser Lys Asp Asp Leu Leu Arg Ile
                 65                  70                  75
Ser Ile Pro Leu Asp Ser Asn Leu Arg Pro Gln Lys Cys Gln Arg
                 80                  85                  90
Phe Ile His Pro Gln Trp Gln Leu Leu His Leu Asn Gly Thr Phe
                 95                 100                 105
Pro Asn Thr Asn Glu Pro Asp Thr Glu Pro Cys Val Asp Gly Trp
                110                 115                 120
Val Tyr Asp Arg Ser Ser Phe Leu Ser Thr Ile Val Thr Glu Trp
                125                 130                 135
Asp Leu Val Cys Glu Ser Gln Ser Leu Lys Ser Met Val Gln Ser
                140                 145                 150
Leu Phe Met Ala Gly Ser Leu Leu Gly Gly Leu Ile Tyr Gly His
                155                 160                 165
Leu Ser Asp Arg Phe Gly Arg Lys Phe Val Leu Arg Trp Ser Tyr
                170                 175                 180
Leu Gln Leu Ala Ile Val Gly Thr Cys Ala Ala Phe Ala Pro Thr
                185                 190                 195
Ile Leu Val Tyr Cys Ser Leu Arg Phe Leu Ala Gly Ala Ala Thr
                200                 205                 210
Phe Ser Ile Ile Val Asn Thr Val Leu Leu Ile Val Glu Trp Ile
                215                 220                 225
Thr His Gln Phe Cys Ala Met Ala Leu Thr Leu Thr Leu Cys Ala
                230                 235                 240
Ala Ser Ile Gly His Ile Thr Leu Gly Ser Leu Ala Phe Val Ile
                245                 250                 255
```

```
Arg Asp Gln Cys Ile Leu Gln Leu Val Met Ser Ala Pro Cys Phe
                260                 265                 270
Val Phe Phe Leu Phe Ser Arg Trp Leu Ala Glu Ser Ala Arg Trp
                275                 280                 285
Leu Ile Ile Asn Asn Lys Pro Glu Glu Gly Leu Lys Glu Leu Arg
                290                 295                 300
Lys Ala Ala His Arg Asn Gly Met Lys Asn Ala Glu Asp Ile Leu
                305                 310                 315
Thr Met Glu Val Leu Lys Ser Thr Met Lys Gln Glu Leu Glu Ala
                320                 325                 330
Ala Gln Lys Lys His Ser Leu Cys Glu Leu Leu Arg Ile Pro Asn
                335                 340                 345
Ile Cys Lys Arg Ile Cys Phe Leu Ser Phe Val Arg Phe Ala Ser
                350                 355                 360
Thr Ile Pro Phe Trp Gly Leu Thr Leu His Leu Gln His Leu Gly
                365                 370                 375
Asn Asn Val Phe Leu Leu Gln Thr Leu Phe Gly Ala Val Thr Leu
                380                 385                 390
Leu Ala Asn Cys Val Ala Pro Trp Ala Leu Asn His Met Ser Arg
                395                 400                 405
Arg Leu Ser Gln Met Leu Leu Met Phe Leu Leu Ala Thr Cys Leu
                410                 415                 420
Leu Ala Ile Ile Phe Val Pro Gln Glu Met Gln Thr Leu Arg Val
                425                 430                 435
Val Leu Ala Thr Leu Gly Val Gly Ala Ala Ser Leu Gly Ile Thr
                440                 445                 450
Cys Ser Thr Ala Gln Glu Asn Glu Leu Ile Pro Ser Ile Ile Arg
                455                 460                 465
Gly Arg Ala Thr Gly Ile Thr Gly Asn Phe Ala Asn Ile Gly Gly
                470                 475                 480
Ala Leu Ala Ser Leu Met Met Ile Leu Ser Ile Tyr Ser Arg Pro
                485                 490                 495
Leu Pro Trp Ile Ile Tyr Gly Val Phe Ala Ile Leu Ser Gly Leu
                500                 505                 510
Val Val Leu Leu Leu Pro Glu Thr Arg Asn Gln Pro Leu Leu Asp
                515                 520                 525
Ser Ile Gln Asp Val Glu Asn Glu Gly Val Asn Ser Leu Ala Ala
                530                 535                 540
Pro Gln Arg Ser Ser Val Leu
                545
```

<210> 31
<211> 988
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473473CD1

<400> 31

```
Met Pro Gly Gly Lys Arg Gly Leu Val Ala Pro Gln Asn Thr Phe
  1               5                  10                  15
Leu Glu Asn Ile Val Arg Arg Ser Ser Glu Ser Ser Phe Leu Leu
                20                  25                  30
Gly Asn Ala Gln Ile Val Asp Trp Pro Val Val Tyr Ser Asn Asp
                35                  40                  45
Gly Phe Cys Lys Leu Ser Gly Tyr His Arg Ala Asp Val Met Gln
                50                  55                  60
Lys Ser Ser Thr Cys Ser Phe Met Tyr Gly Glu Leu Thr Asp Lys
                65                  70                  75
Lys Thr Ile Glu Lys Val Arg Gln Thr Phe Asp Asn Tyr Glu Ser
                80                  85                  90
Asn Cys Phe Glu Val Leu Leu Tyr Lys Lys Asn Arg Thr Pro Val
                95                 100                 105
Trp Phe Tyr Met Gln Ile Ala Pro Ile Arg Asn Glu His Glu Lys
               110                 115                 120
```

```
Val Val Leu Phe Leu Cys Thr Phe Lys Asp Ile Thr Leu Phe Lys
125                 130                 135
Gln Pro Ile Glu Asp Asp Ser Thr Lys Gly Trp Thr Lys Phe Ala
140                 145                 150
Arg Leu Thr Arg Ala Leu Thr Asn Ser Arg Ser Val Leu Gln Gln
155                 160                 165
Leu Thr Pro Met Asn Lys Thr Glu Val Val His Lys His Ser Arg
170                 175                 180
Leu Ala Glu Val Leu Gln Leu Gly Ser Asp Ile Leu Pro Gln Tyr
185                 190                 195
Lys Gln Glu Ala Pro Lys Thr Pro Pro His Ile Ile Leu His Tyr
200                 205                 210
Cys Ala Phe Lys Thr Thr Trp Asp Trp Val Ile Leu Ile Leu Thr
215                 220                 225
Phe Tyr Thr Ala Ile Met Val Pro Tyr Asn Val Ser Phe Lys Thr
230                 235                 240
Lys Gln Asn Asn Ile Ala Trp Leu Val Leu Asp Ser Val Val Asp
245                 250                 255
Val Ile Phe Leu Val Asp Ile Val Leu Asn Phe His Thr Thr Phe
260                 265                 270
Val Gly Pro Gly Gly Glu Val Ile Ser Asp Pro Lys Leu Ile Arg
275                 280                 285
Met Asn Tyr Leu Lys Thr Trp Phe Val Ile Asp Leu Leu Ser Cys
290                 295                 300
Leu Pro Tyr Asp Ile Ile Asn Ala Phe Glu Asn Val Asp Glu Gly
305                 310                 315
Ile Ser Ser Leu Phe Ser Ser Leu Lys Val Val Arg Leu Leu Arg
320                 325                 330
Leu Gly Arg Val Ala Arg Lys Leu Asp His Tyr Leu Glu Tyr Gly
335                 340                 345
Ala Ala Val Leu Val Leu Leu Val Cys Val Phe Gly Leu Val Ala
350                 355                 360
His Trp Leu Ala Cys Ile Trp Tyr Ser Ile Gly Asp Tyr Glu Val
365                 370                 375
Ile Asp Glu Val Thr Asn Thr Ile Gln Ile Asp Ser Trp Leu Tyr
380                 385                 390
Gln Leu Ala Leu Ser Ile Gly Thr Pro Tyr Arg Tyr Asn Thr Ser
395                 400                 405
Ala Gly Ile Trp Glu Gly Gly Pro Ser Lys Asp Ser Leu Tyr Val
410                 415                 420
Ser Ser Leu Tyr Phe Thr Met Thr Ser Leu Thr Thr Ile Gly Phe
425                 430                 435
Gly Asn Ile Ala Pro Thr Thr Asp Val Glu Lys Met Phe Ser Val
440                 445                 450
Ala Met Met Met Val Gly Ala Leu Leu Tyr Ala Thr Ile Phe Gly
455                 460                 465
Asn Val Thr Thr Ile Phe Gln Gln Met Tyr Ala Asn Thr Asn Arg
470                 475                 480
Tyr His Glu Met Leu Asn Asn Val Arg Asp Phe Leu Lys Leu Tyr
485                 490                 495
Gln Val Pro Lys Gly Leu Ser Glu Arg Val Met Asp Tyr Ile Val
500                 505                 510
Ser Thr Trp Ser Met Ser Lys Gly Ile Asp Thr Glu Lys Val Leu
515                 520                 525
Ser Ile Cys Pro Lys Asp Met Arg Ala Asp Ile Cys Val His Leu
530                 535                 540
Asn Arg Lys Val Phe Asn Glu His Pro Ala Phe Arg Leu Ala Ser
545                 550                 555
Asp Gly Cys Leu Arg Ala Leu Ala Val Glu Phe Gln Thr Ile His
560                 565                 570
Cys Ala Pro Gly Asp Leu Ile Tyr His Ala Gly Glu Ser Val Asp
575                 580                 585
Ala Leu Cys Phe Val Val Ser Gly Ser Leu Glu Val Ile Gln Asp
590                 595                 600
Asp Glu Val Val Ala Ile Leu Gly Lys Gly Asp Val Phe Gly Asp
605                 610                 615
Ile Phe Trp Lys Glu Thr Thr Leu Ala His Ala Cys Ala Asn Val
```

620 625 630
Arg Ala Leu Thr Tyr Cys Asp Leu His Ile Ile Lys Arg Glu Ala
635 640 645
Leu Leu Lys Val Leu Asp Phe Tyr Thr Ala Phe Ala Asn Ser Phe
650 655 660
Ser Arg Asn Leu Thr Leu Thr Cys Asn Leu Arg Lys Arg Ile Ile
665 670 675
Phe Arg Lys Ile Ser Asp Val Lys Lys Glu Glu Glu Glu Arg Leu
680 685 690
Arg Gln Lys Asn Glu Val Thr Leu Ser Ile Pro Val Asp His Pro
695 700 705
Val Arg Lys Leu Phe Gln Lys Phe Lys Gln Gln Lys Glu Leu Arg
710 715 720
Asn Gln Gly Ser Thr Gln Gly Asp Pro Glu Arg Asn Gln Leu Gln
725 730 735
Val Glu Ser Arg Ser Leu Gln Asn Gly Ala Ser Ile Thr Gly Thr
740 745 750
Ser Val Val Thr Val Ser Gln Ile Thr Pro Ile Gln Thr Ser Leu
755 760 765
Ala Tyr Val Lys Thr Ser Glu Ser Leu Lys Gln Asn Asn Arg Asp
770 775 780
Ala Met Glu Leu Lys Pro Asn Gly Gly Ala Asp Gln Lys Cys Leu
785 790 795
Lys Val Asn Ser Pro Ile Arg Met Lys Asn Gly Asn Gly Lys Gly
800 805 810
Trp Leu Arg Leu Lys Asn Asn Met Gly Ala His Glu Glu Lys Lys
815 820 825
Glu Asp Trp Asn Asn Val Thr Lys Ala Glu Ser Met Gly Leu Leu
830 835 840
Ser Glu Asp Pro Lys Ser Ser Asp Ser Glu Asn Ser Val Thr Lys
845 850 855
Asn Pro Leu Arg Lys Thr Asp Ser Cys Asp Ser Gly Ile Thr Lys
860 865 870
Ser Asp Leu Arg Leu Asp Lys Ala Gly Glu Ala Arg Ser Pro Leu
875 880 885
Glu His Ser Pro Ile Gln Ala Asp Ala Lys His Pro Phe Tyr Pro
890 895 900
Ile Pro Glu Gln Ala Leu Gln Thr Thr Leu Gln Glu Val Lys His
905 910 915
Glu Leu Lys Glu Asp Ile Gln Leu Leu Ser Cys Arg Met Thr Ala
920 925 930
Leu Glu Lys Gln Val Ala Glu Ile Leu Lys Ile Leu Ser Glu Lys
935 940 945
Ser Val Pro Gln Ala Ser Ser Pro Lys Ser Gln Met Pro Leu Gln
950 955 960
Val Pro Pro Gln Ile Pro Cys Gln Asp Ile Phe Ser Val Ser Arg
965 970 975
Pro Glu Ser Pro Glu Ser Asp Lys Asp Glu Ile His Phe
980 985

<210> 32
<211> 533
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477725CD1

<400> 32
Met Ala Phe Glu Glu Leu Leu Ser Gln Val Gly Gly Leu Gly Arg
1 5 10 15
Phe Gln Met Leu His Leu Val Phe Ile Leu Pro Ser Leu Met Leu
20 25 30
Leu Ile Pro His Ile Leu Leu Glu Asn Phe Ala Ala Ala Ile Pro
35 40 45
Gly His Arg Cys Trp Val His Met Leu Asp Asn Asn Thr Gly Ser

```
                    50                  55                  60
Gly Asn Glu Thr Gly Ile Leu Ser Glu Asp Ala Leu Leu Arg Ile
                65                  70                  75
Ser Ile Pro Leu Asp Ser Asn Leu Arg Pro Glu Lys Cys Arg Arg
                80                  85                  90
Phe Val His Pro Gln Trp Gln Leu Leu His Leu Asn Gly Thr Ile
                95                 100                 105
His Ser Thr Ser Glu Ala Asp Thr Glu Pro Cys Val Asp Gly Trp
               110                 115                 120
Val Tyr Asp Gln Ser Tyr Phe Pro Ser Thr Ile Val Thr Lys Trp
               125                 130                 135
Asp Leu Val Cys Asp Tyr Gln Ser Leu Lys Ser Val Val Gln Phe
               140                 145                 150
Leu Leu Leu Thr Gly Met Leu Val Gly Gly Ile Ile Gly Gly His
               155                 160                 165
Val Ser Asp Arg Phe Gly Arg Arg Phe Ile Leu Arg Trp Cys Leu
               170                 175                 180
Leu Gln Leu Ala Ile Thr Asp Thr Cys Ala Ala Phe Ala Pro Thr
               185                 190                 195
Phe Pro Val Tyr Cys Val Leu Arg Phe Leu Ala Gly Phe Ser Ser
               200                 205                 210
Met Ile Ile Ile Ser Asn Asn Ser Leu Pro Ile Thr Glu Trp Ile
               215                 220                 225
Arg Pro Asn Ser Lys Ala Leu Val Val Ile Leu Ser Ser Gly Ala
               230                 235                 240
Leu Ser Ile Gly Gln Ile Ile Leu Gly Gly Leu Ala Tyr Val Phe
               245                 250                 255
Arg Asp Trp Gln Thr Leu His Val Val Ala Ser Val Pro Phe Phe
               260                 265                 270
Val Phe Phe Leu Leu Ser Arg Trp Leu Val Glu Ser Ala Arg Trp
               275                 280                 285
Leu Ile Ile Thr Asn Lys Leu Asp Glu Gly Leu Lys Ala Leu Arg
               290                 295                 300
Lys Val Ala Arg Thr Asn Gly Ile Lys Asn Ala Glu Glu Thr Leu
               305                 310                 315
Asn Ile Glu Val Val Arg Ser Thr Met Gln Glu Glu Leu Asp Ala
               320                 325                 330
Ala Gln Thr Lys Thr Thr Val Cys Asp Leu Phe Arg Asn Pro Ser
               335                 340                 345
Met Arg Lys Arg Ile Cys Ile Leu Val Phe Leu Arg Phe Ala Asn
               350                 355                 360
Thr Ile Pro Phe Tyr Gly Thr Met Val Asn Leu Gln His Val Gly
               365                 370                 375
Ser Asn Ile Phe Leu Leu Gln Val Leu Tyr Gly Ala Val Ala Leu
               380                 385                 390
Ile Val Arg Cys Leu Ala Leu Leu Thr Leu Asn His Met Gly Arg
               395                 400                 405
Arg Ile Ser Gln Ile Leu Phe Met Phe Leu Val Gly Leu Ser Ile
               410                 415                 420
Leu Ala Asn Thr Phe Val Pro Lys Glu Met Gln Thr Leu Arg Val
               425                 430                 435
Ala Leu Ala Cys Leu Gly Ile Gly Cys Ser Ala Ala Thr Phe Ser
               440                 445                 450
Ser Val Ala Val His Phe Ile Glu Leu Ile Pro Thr Val Leu Arg
               455                 460                 465
Ala Arg Ala Ser Gly Ile Asp Leu Thr Ala Ser Arg Ile Gly Ala
               470                 475                 480
Ala Leu Ala Pro Leu Leu Met Thr Leu Thr Val Phe Phe Thr Thr
               485                 490                 495
Leu Pro Trp Ile Ile Tyr Gly Ile Phe Pro Ile Ile Gly Gly Leu
               500                 505                 510
Ile Val Phe Leu Leu Pro Glu Thr Lys Asn Leu Pro Leu Pro Asp
               515                 520                 525
Thr Ile Lys Asp Val Glu Asn Gln
               530
```

<210> 33

<211> 1775
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3474673CB1

<400> 33
atttcaggaa atgtgagggg gctctgggcc ccttccctca gcgcctgcgg tcacccagca 60
gcttcctcct tctccctggc ctaggcctag caggtgggca ccccgcacac atttgaggcg 120
gggccagatg cccacagttc agagcctctt tttgtcccgg ggattggatc ccagggctgg 180
gtggggccag gctgtcccat tccccaacac tcctcctccc cggcgaaacc gggcaccagc 240
aggcgtttgc gagaggagat acgagctgga cgcctggccc ttccctccca ccgggtccta 300
gtccaccgct cccggcgccg gctccccgcc tctcccgcta tgtaccgacc gcgagcccgg 360
gcggctcccg agggcagggt ccggggctgc gcggtgccca gcaccgtgct cctgctgctc 420
gcctacctgg cttacctggc gctgggcacc ggcgtgttct ggacgctgga gggccgcgcg 480
gcgcaggact ccagccgcag cttcagcgc gacaagtggg agctgttgca gaacttcacg 540
tgtctggacc gcccggcgct ggactcgctg atccgggatg tcgtccaagc atacaaaaac 600
ggagccagcc tcctcagcaa caccaccagc atggggcgct gggagctcgt gggctccttc 660
ttcttttctg tgtccaccat caccaccatt ggctatggca acctgagccc caacacgatg 720
gctgcccgcc tcttctgcat cttctttgcc cttgtgggga tcccactcaa cctcgtggtg 780
ctcaaccgac tggggcatct catgcagcag ggagtaaacc actgggccag caggctgggg 840
ggcacctggc aggatcctga caaggcgcgg tggctggcgg gctctggcgc cctcctctcg 900
ggcctcctgc tcttcctgct gctgccaccg ctgctcttct cccacatgga gggctggagc 960
tacacagagg gcttctactt cgccttcatc accctcagca ccgtgggctt cggcgactac 1020
gtgattggaa tgaacccctc ccagaggtac ccactgtggt acaagaacat ggtgtcctg 1080
tggatcctct ttgggatggc atggctggcc ttgatcatca aactcatcct ctcccagctg 1140
gagacgccag ggagggtatg ttcctgctgc caccacagct ctaaggaaga cttcaagtcc 1200
caaagctgga gacagggacc tgaccgggag ccagagtccc actccccaca gcaaggatgc 1260
tatccagagg gacccatggg aatcatacag catctggaac cttctgctca cgctgcaggc 1320
tgtggcaagg acagctagtt atactccatt ctttggtcgt cgtcctcggt agcaagaccc 1380
ctgattttaa gctttgcaca tgtccaccca aactaaagac tacattttcc atccacccta 1440
gaggctgggt gcagctatat gattaattct gcccaatagg gtatacagag acatgtcctg 1500
ggtgacatgg gatgtgactt tcgggtgtcg gggcagcatg cccttctccc ccacttcctt 1560
actttagcgg gctgcaatgc cgccgatatg atggctggga gctctggcag ccatacggca 1620
ccatgaagta gcggcaatgt ttgagcggca caataagata ggaagagtct ggatctctga 1680
tgatcacaga gccatcctaa caaacggaat atcacccgac ctcctttatg tgagagagaa 1740
ataaacatct tatgtaaaat accaaaaaaa aaaaa 1775

<210> 34
<211> 1545
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4588877CB1

<400> 34
aatgagggcc ctgggggttg ggcccaggag tggggctgtg gtggtgagtg gacagggctg 60
ggctggaaat gtcccctgag tgccccctct cacctcaggc tatggcaacc tgagcccaa 120
cacgatggct gcccgcctct tctgcatctt ctttgccctt gtggggatcc cactcaacct 180
cgtggtgctc aaccgactgg ggcatctcat gcagcaggga gtaaaccact gggccagcag 240
gctgggggggc acctggcagg tgagggggct gctggacggg gtggggatgg gtcacttcta 300
gaatgagggg ctgtggtggg aattggggtt actaatgaca agaggtggga gcaagtgtta 360
ctggtgaggt tgtgttggga ttgggggtca ctgctcagaa tagggtcctt agtgaaaaag 420
ggcattaatg gtggagatgg ggtgggactg ggcagacagg aaggacatga ggcacaggct 480
ccaggcaggg aacctggaga acacagacca ggtgaagagc ccccttctta ctggggacag 540
ctctggcctg cctccagctc cctcggctcc cacgcatggg gtgaaggcct caggaggcct 600
ggggacaata ttgcacccac aggatcctga caaggcgcgg tggctggcgg gctctggcgc 660
cctcctctcg ggcctcctgc tcttcctgct gctgccaccg ctgctcttct cccacatgga 720
gggctggagc tacacagagg gcttctactt cgccttcatc accctcagca ccgtgggctt 780
cggcgactac gtgattggaa tgaacccctc ccagaggtac ccactgtggt acaagaacat 840
ggtgtccctg tggatcctct ttgggatggc atggctggcc ttgatcatca aactcatcct 900
ctcccagctg gagacgccag ggagggtatg ttcctgctgc caccacagct ctaaggaaga 960
cttcaagtcc caaagctgga gacagggacc tgaccgggag ccagagtccc actccccaca 1020

gcaaggatgc tatccagagg gacccatggg aatcatacag catctggaac attctgctca 1080
cgctgcaggc tgtggcaagg acagctagtt atactccatt ctttggtcgt cgtcctcggt 1140
agcaagaccc ctgatttttaa gctttgcaca tgtccaccca aactaaagat tacatttttcc 1200
atccacccta gaggctgggt gcagctatat gattaattct gcccaatagg gtatacagag 1260
acatgtcctg ggtgacatgg gatgtgactt tcgggtgtcg gggcagcatg cccttctccc 1320
ccacttcctt actttagcgg gctgcaatgc cgccgatatg atggctggga gctctggcag 1380
ccatacggca ccatgaagta gcggcaatgt ttgagcggca caataagata ggaagagtct 1440
ggatctctga tgatcacaga gccatcctaa caaacggaat atcacccgac ctcctttatg 1500
tgagagagaa ataaacatct tatgtaaaat accaaaaaaa aaaaa 1545

<210> 35
<211> 1941
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472214CB1

<400> 35
atggcggaga aggcgctgga ggccgtgggc tgtggactag ggccgggggc tgtggccatg 60
gccgtgacgc tggaggacgg ggcggaaccc cctgtgctga ccacgcacct gaagaaggtg 120
gagaaccaca tcactgaagc ccagcgcttc tcccacctac ccaagcgctc agccgtggac 180
atcgagttcg tggagctgtc ctattccgtg cgggaggggc cctgctggcg caaaaggggt 240
tataagaccc ttctcaagtg cctctcaggt aaattctgcc gccgggagct gattggcatc 300
atgggccccc caggggctgg caagtctaca ttcatgaaca tcttggcagg atacagggag 360
tctggaatga aggggcagat cctggttaat ggaaggccac gggagctgag gaccttccgc 420
aagatgtcct gctacatcat gcaagatgac atgctgctgc cgcaccttac ggtgttggaa 480
gccatgatgg tgtctgctaa cctgaatctt actgagaatc ccgatgtgaa aaacgatctc 540
gtgacagaga tcctgacggc actgggcctg atgtcgtgct cccacacgag gacagccctg 600
ctctctggcg ggcagaggaa gcgtctggcc atcgccctgg agctggtcaa caacccgcct 660
gtcatgttct ttgatgagcc caccagtggt ctggatagcg cctcttgttt ccaagtggtg 720
tccctcatga agtccctggc acaggggggc cgtaccatca tctgcaccat ccaccagccc 780
agtgccaagc tctttgagat gtttgacaag ctctacatcc tgagccaggg tcagtgcatc 840
ttcaaaggcg tggtcaccaa cctgatcccc tatctaaagg gactcggctt gcattgcccc 900
acctaccaca acccggctga cttcgtcatc gaggtggcct ctggcgagta tggagacctg 960
aaccccatgt tgttcagggc tgtgcagaat gggctgtgcg ctatggctga gaagaagagc 1020
agccctgaga agaacgaggt ccctgcccca tgccctcctt gtcctccgga agtggatccc 1080
attgaaagcc acacctttgc caccagcacc ctcacacagt tctgcatcct cttcaagagg 1140
accttcctgt ccatcctcag ggacacggtg ctgacccacc tacggttcat gtcccacgtg 1200
gttattggcg tgctcatcgg cctcctctac ctgcatattg gcgacgatgc cagcaaggtc 1260
ttcaacaaca ccggctgcct cttcttctcc atgctgttcc tcatgttcgc cgccctcatg 1320
ccaactgtgc tcaccgtccc cttagagatg gcggtcttca tgagggagca cctcaactac 1380
tggtacagcc tcaaagcgta ttacctggcc aagaccatgg ctgacgtgcc ctttcaggtg 1440
gtgtgtccgg tggtctactg cagcattgtg tactggatga cgggccagcc cgctgagacc 1500
agccgcttcc tgctcttctc agccctggcc accgccaccg ccttggtggc ccaatctttg 1560
gggctgctga tcggagctgc ttccaactcc ctacaggtgg gcacttttgt gggcccagtt 1620
accgccatcc ctgtcctctt gttctccggc ttctttgtca gcttcaagac catccccact 1680
tacctgcaat ggagctccta tctctcctat gtcaggtatg gctttgaggg tgtgatcctg 1740
acgatctatg gcatggagcg aggagacctg acatgtttag aggaacgctg cccgttccgg 1800
gagccacaga gcatcctccg agcgctggat gtggaggatg ccaagctcta catggacttc 1860
ctggtcttgg gcatcttctt cctagccctg cggctgctgg cctaccttgt gctgcgttac 1920
cgggtcaagt cagagagata g 1941

<210> 36
<211> 4971
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473053CB1

<400> 36
caaagtagcg ggccgaggcc cgggggagcg gggccgcagc tgggggggcg ggagcccgtg 60
gggagccgag ccgagcgccc cccgccccag cccccggcat gggcagtacg gggccgccgg 120
ggcgggcgcc gagcgctgag cgctgagggt ctcccatggg attgctggga tcttgctggg 180

tgagatggca gtgtgtgcaa aaaagcgccc cccagaagaa gaaaggaggg cgcgggctaa 240
tgaccgagaa tacaatgaga aattccagta tgcgagtaac tgcatcaaga cctccaagta 300
caatattctc accttcctgc ctgtcaacct ctttgagcag ttccaggaag ttgccaacac 360
ttacttcctg ttcctcctca ttctgcagtt gatcccccag atctcttccc tgtcctggtt 420
caccaccatt gtgcctttgg ttcttgtcct caccatcaca gctgttaaag atgccactga 480
tgactatttc cgccacaaga gcgataacca ggtgaataac cgccagtctc aggtgctgat 540
caacggaatc ctccagcagg agcagtggat gaatgtctgt gttggtgata ttatcaagct 600
agaaaataac cagtttgtgg cggcggatct cctcctcctt tccagcagtg agccccatgg 660
gctgtgttac atagagacag cagaacttga tggcgagacc aacatgaaag tacgtcaggc 720
gattccagtc acctcagaat tgggagacat cagtaagctt gccaagtttg acggtgaagt 780
gatctgtgaa cctcccaaca acaaactgga caaattcagc ggaaccctct actggaagga 840
aaataagttc cctctgagca accagaacat gctgctgcgg ggctgtgtgc tgcgaaacac 900
cgagtggtgc ttcgggctgg tcatctttgc aggtcccgac actaagctga tgcaaaacag 960
cggcagaaca aagttcaaaa gaacgagtat cgatcgccta atgaataccc tggtgctctg 1020
gatttttgga ttcctggttt gcatgggggt gatcctggcc attggcaatg ccatctggga 1080
gcacgaggtg gggatgcgtt tccaggtcta cctgccgtgg gatgaggcag tggacagtgc 1140
cttcttctct ggcttcctct ccttctggtc ctacatcatc atcctcaaca ccgttgtgcc 1200
catttcactc tatgtcagtg tggaggtcat ccgtctgggc cacagctact tcatcaactg 1260
ggataagaag atgttctgca tgaagaagcg gacgcctgca gaagcccgca ccaccaccct 1320
aaacgaggag ctgggccagg tggagtacat cttctccgac aagacgggca ccctcaccca 1380
gaacatcatg gttttcaaca agtgctccat caatggccac agctatggtg atgtgtttga 1440
cgtcctggga cacaaagctg aattgggaga gaggcctgaa cctgttgact tctccttcaa 1500
tcctctggct gacaagaagt tcttattttg ggaccccagc ctgctggagg ctgtcaagat 1560
cggggacccc cacacgcatg agttcttccg cctcctttcc ctgtgtcata ctgtcatgtc 1620
agaagaaaag aacgaaggag agctgtacta caaagctcag tccccagatg agggggccct 1680
ggtcaccgca gccaggaact ttggttttgt tttccgctct cgcaccccca aaacaatcac 1740
cgtccatgag atgggcacag ccatcaccta ccagctgctg gccatcctgg acttcaacaa 1800
catccgcaag cggatgtcgg tcatagtgcg gaatccagag gggaagatcc gactctactg 1860
caaaggggct gacactatcc tactggacag actgcaccac tccactcaag agctgctcaa 1920
caccaccatg gaccacctta atgagtacgc aggggaaggg ctgaggaccc tggtgctggc 1980
ctacaaggat ctggatgaag agtactatga ggagtgggct gagcgacgcc tccaggccag 2040
cctggcccag gacagccggg aggacaggct ggctagcatc tatgaggagg ttgagaacaa 2100
catgatgctg ctgggtgcaa cggccattga ggacaaactt cagcaagggg ttccagagac 2160
cattgccctc ctgacactgg ccaacatcaa gatttgggtg ctaaccggag acaagcaaga 2220
gacggctgtg aacatcggct attcctgcaa gatgctgacg gatgacatga ctgaggtttt 2280
catagtcact ggccatactg tcctggaggt gcgggaggag ctcaggaaag cccgggagaa 2340
gatgatggac tcatcccgct ctgtaggcaa cggcttcacc tatcaggaca agctttcttc 2400
ttccaagcta acttctgtcc tggaggccgt tgctggggag tacgccctgg tcataaatgg 2460
tcacagcctg gcccacgcac tggaggcaga catggagctg gagtttctgg agacagcgtg 2520
tgcctgcaaa gctgtcatct gctgccgggt gacccccttg cagaaggcac aggtggtaga 2580
actggtcaag aagtacaaga aggctgtgac gcttgccatt ggagacggag ccaatgatgt 2640
cagcatgatc aaaacggctc acattggtgt ggggatcagt gggcaggaag ggatccaggc 2700
tgtcttggcc tccgattact ccttctccca gttcaagttc ctgcagcgcc tcctgctggt 2760
gcatgggcgc tggtcctacc tgcgaatgtg caagtttctt tgctatttct tctacaaaaa 2820
ctttgctttc accatggtcc acttctggtt tggcttcttc tgtggcttct cagcccagac 2880
cgtctatgac cagtatttca tcaccctgta taacatcgtg tacacctccc tgccagtcct 2940
ggctatgggg gtctttgatc aggatgtccc cgagcagcgg agcatggagt accctaagct 3000
gtatgagccg ggccagctga accttctctt caacaagcgg gagttcttca tctgcatcgc 3060
ccagggcatc tacacctccg tgctcatgtt cttcattccc tatggggtgt ttgctgatgc 3120
cacccgggat gatggcactc agctggctga ctaccagtcc tttgcagtca ctgtggccac 3180
atccttggtc attgtggtta gcgtgcagat tgggctcgac acaggctact ggacggccat 3240
caaccacttc ttcatctggg gaagccttgc tgtttacttt gccatcctct ttgccatgca 3300
cagcaatggg ctcttcgaca tgtttcccaa ccagttccgg tttgtgggga atgcccagaa 3360
caccttggcc cagcccacgg tgtggctgac cattgtgctc accacagtcg tctgcatcat 3420
gcccgtggtt gccttccgat tcctcaggct caacctgaag ccggatctct ccgacacggt 3480
ccgctacaca cagctcgtga ggaagaagca gaaggcccag caccgctgca tgcggcgggt 3540
tggccgcact ggctcccggc gctccggcta tgccttctcc catcaggagg gcttcgggga 3600
gctcatcatg tctggcaaga acatgcggct gagctctctc gcgctctcca gcttcaccac 3660
ccgctccagc tccagctgga ttgagagcct gcgcaggaag aagagtgaca gtgccagtag 3720
ccccagtggc ggtgccgaca agccctcaa gggctgaagg ccgaggatgg atgccctgtg 3780
ccagtgacca gagcacccag ggctggccag tcactgaggg aacagcgtct cggaactgct 3840
ggtcctcatt ccttgcttcc cgtcccccg gtagactctg tcctgctggt cccaccacac 3900
atggctggga catctgttcc cagctgtagg cccttccacc agctggggag ctagagggag 3960
caggcccaag ggcagagcag aggctgaggc acggggagcc agccccactc ggggaccaga 4020
agtggaacca aaaacaagaa aaaactgtga gagattgtgt ctgcccctgc cctgcctggg 4080
acccacaggg agactataat ctccttattt ttttactcct actccccaga ggggccctag 4140
tgcctctgtt cctgaattac ataagaatgt accatgccgg gaagccagag acctgcaggg 4200

gcctcggccc ctcacatcgt gtatgtctct ccttgatttg tgttgtgtcc agtttggttt 4260
tgtcttttt tatttggcaa gtggaggagg cttttatgtg acttttatgt tgtggttggt 4320
gtcttaactc tcctgggaaa aggaggctgg cacacactgg gatgccgcag cctggccggc 4380
tgtggggtgg tttgggagga tccatgtcgg ctctgcctgc agtgaccagt gctctgtggg 4440
gcagaggagc tgaccaggga gggaggtacc catgagcaga gggtagtggg agagtgtaaa 4500
ggagggtttg gtcctgtctg cttcctcacc ttgagagtaa agtgctgccc tctgccccca 4560
acacacacac atatcaattc ctggattcct tagtcctgct ggccttgggc tggagcctag 4620
gaaagtggcc cccaaatcct tagtgagcta aagctgggtc tgaaatttgg tcagtgggga 4680
ggggtagttt tcttttcttt tttcttttte ttttttttct ttttttttg agatggagtc 4740
tcactcttgt cacctaggca agagtgcaat ggcacaatct cagctcactg caacctccac 4800
ctcctgggtt caagcgattc tcctgcctcc ccggacccaa ccactggact taatctcact 4860
ttcttaaatt cttctattct cagacacggg tctagtacca ttccttcctc ttagccccag 4920
ggagcaaatt aaagaggtta cgagttaaaa tcctaaaaaa aaaaaaaaaa a 4971

<210> 37
<211> 1404
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473347CB1

<400> 37
atggtcctgg ctttccagtt agtctccttc acctacatct ggatcatatt gaaaccaaat 60
gtttgtgctg cttctaacat caagatgaca caccagcggt gctcctcttc aatgaaacaa 120
acctgcaaac aagaaactag aatgaagaaa gatgacagta ccaaagcgcg gcctcagaaa 180
tatgagcaac ttctccatat agaggacaac gatttcgcaa tgagacctgg atttggaggg 240
tctccagtgc cagtaggtat agatgtccat gttgaaagca ttgacagcat ttcagagact 300
aacatggact ttacaatgac tttttatctc aggcattact ggaaagacga gaggctctcc 360
tttcctagca cagcaaacaa aagcatgaca tttgatcata gattgaccag aaagatctgg 420
gtgcctgata tctttttttgt ccactctaaa agatccttca tccatgatac aactatggag 480
aatatcatgc tgcgcgtaca ccctgatgga aacgtcctcc taagtctcag gataacggtt 540
tcggccatgt gctttatgga tttcagcagg tttcctcttg acactcaaaa ttgttctctt 600
gaactggaaa gctatgccta caatgaggat gacctaatgc tatactggaa acacggaaac 660
aagtccttaa atactgaaga acatatgtcc ctttctcagt tcttcattga agacttcagt 720
gcatctagtg gattagcttt ctatagcagc acaggctggt acaataggct tttcatcatc 780
tctgtgctaa ggaggcatgt tttcttcttt gtgctgccaa cctattaccc agccatattg 840
atggtgatgc tttcatgggt ttcattttgg attgaccgaa gagctgttcc tgcaagagtt 900
tccctgggaa tcaccacagt gctgaccatg tccacaatca tcactgctgt gagcgcctcc 960
atgccccagg tgtcctacct caaggctgtg gatgtgtacc tgtgggtcag ctccctcttt 1020
gtgttcctgt cagtcattga gtatgcagct gtgaactacc tcaccacagt ggaagagcgg 1080
aaacaattca agaagacagg aaagatttct aggatgtaca atattgatgc agttcaagct 1140
atggcctttg atggttgtta ccatgacagc gagattgaca tggaccagac ttccctctct 1200
ctaaactcag aagacttcat gagaagaaaa tcgatatgca gccccagcac cgattcatct 1260
cggataaaga gaagaaaatc cctaggagga catgttggta gaatcattct ggaaaacaac 1320
catgtcattg acacctattc taggatttta ttccccattg tgtatatttt atttaatttg 1380
ttttactggg gtgtatatgt atga 1404

<210> 38
<211> 4048
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474240CB1

<400> 38
cttccatccc ccctcagcca ttccttactg ctctgggcaa ccgccaggtt aagcccattt 60
gcactgggaa attggcgctg tttgggagaa gagaaacaga tcgattgccc ttgtgactcc 120
ccgccccctt cccatcccca cccccaccgc tctctccctc tttccctccc ccgccacctc 180
ccctcacccc gcctccttcc cgttccccac ccccaaaccc tctcacccgc ggcagtccgg 240
tgcgaggccc cctccggaag gtgaggggaa tggattggac tccggtggag aaagcgggtg 300
tctagaagtg gtgctaatgg gaagagaatt ctggtttcaa aagaggatgc tctgccacaa 360
agagcggctc gcgcgctggc ctgggctcta gccgaggaga gatcccggga ggactccaga 420
gctccggggg agcgctcctc ggaagaccgg ggccaacatg cctgtgcgca gggggcatgt 480

ggcaccacaa aatacatttc tggggaccat cattcggaaa tttgaagggc aaaataaaaa 540
atttatcatt gcaaatgcca gagtgcagaa ctgtgccatc atttattgca acgatgggtt 600
ctgtgagatg actggtttct ccaggccaga tgtcatgcaa aagccatgca cctgcgactt 660
tctccatgga cccgagacca agaggcatga tattgcccaa attgcccagg cattgctggg 720
gtcagaagag aggaaagtgg aggtcaccta ctatcacaaa aatgggtcca cttttatttg 780
taacactcac ataattccag tgaaaaacca agagggcgtg gctatgatgt tcatcattaa 840
ttttgaatat gtgacggata atgaaaacgc tgctacccca gagagggtaa acccaatatt 900
accaatcaaa actgtaaacc ggaaattttt tgggttcaaa ttccctggtc tgagagttct 960
cacttacaga aagcagtcct taccacaaga agaccccgat gtggtggtca tcgattcatc 1020
taaacacagt gatgattcag tagccatgaa gcattttaag tctcctacaa aagaaagctg 1080
cagcccctct gaagcagatg acacaaaagc tttgatacag cccagcaaat gttctccctt 1140
ggtgaatata tccggacctc ttgaccattc ctctcccaaa aggcaatggg accgactcta 1200
ccctgacatg ctgcagtcaa gttcccagct gtcccattcc agatcaaggg aaagcttatg 1260
tagtatacgg agagcatctt cggtccatga tatagaagga ttcggcgtcc accccaagaa 1320
catatttaga gaccgacatg ccagcgaaga caatggtcgc aatgtcaaag ggccttttaa 1380
tcatatcaag tcaagcctcc tgggatccac atcagattca aacctcaaca aatacagcac 1440
cattaacaag attccacagc tcactctgaa ttttcagag gtcaaaactg agaaaaagaa 1500
ttcatcacct ccttcttcag ataaaaccat tattgcaccc aaggttaaag atcgaacaca 1560
caatgtgact gagaaagtga cccaggttct ctctttagga gcagatgtcc tacctgaata 1620
caaactgcag acaccacgca tcaacaagtt tacgatattg cactacagcc ctttcaaggc 1680
agtctgggac tggcttatcc tgctgttggt catatacact gctatattta ctccctactc 1740
tgcagccttc ctcctcaatg acagagaaga acagaaaaga cgagaatgtg gctattcttg 1800
tagccctttg aatgtggtag acttgattgt ggatattatg tttatcatag atattttaat 1860
aaacttcaga acaacatatg taaatcagaa tgaagaagtg gtaagtgatc ccgccaaaat 1920
agcaatacac tacttcaaag gctggttcct gattgacatg gttgcagcaa ttccttttga 1980
cttgctgatt tttggatcag gttctgatga gacaacaaca ttaattggtc ttttgaagac 2040
tgcccgactc ctccgtcttg tgcgggtggc caggaaactg gatcgatatt cagaatatgg 2100
cgctgctgtt ctaatgctct taatgtgcat ctttgccctg attgctcact ggctggcttg 2160
catttggtat gcgattggga atgtagaaag gccttacctg actgacaaaa tcggatggtt 2220
ggattcctta ggacagcaaa ttgggaaacg ttacaatgac agtgactcaa gttctggacc 2280
atccattaaa gacaaatacg tcacagcact ttattttacc ttcagcagtt taaccagtgt 2340
aggattcggg aatgtgtctc ctaacacgaa ttcggagaaa atcttttcaa tttgtgtcat 2400
gttgattggc tcactaatgt atgcaagcat ttttgggaat gtatctgcaa ttatccaaag 2460
actatactcg ggaactgcca ggtaccacat gcagatgctg cgagtaaaag agttcattcg 2520
ctttcaccaa atccccaacc ctctgaggca acgtcttgaa gaatatttcc agcacgcatg 2580
gacttacacc aatggcattg acatgaacat ggtcctaaag ggtttcccag aatgcttaca 2640
agcagacatt tgtctacatc tcaaccagac attgctgcaa aactgcaaag cctttcgggg 2700
ggcaagtaaa ggttgcctta gagctttggc aatgaagttc aaaaccaccc atgcacctcc 2760
aggagacacc ctcgttcact gtggggatgt cctcactgca ctttatttct tatccagagg 2820
ctccattgaa attctcaaag atgacattgt ggtggctatt ctgggaaaaa atgatatatt 2880
tggagaaatg gttcatcttt atgccaaacc tggaaagtct aatgcagatg taagagccct 2940
cacatactgt gacttgcata agattcagcg agaagacttg ttagaggttt tggatatgta 3000
tcctgagttt tctgatcact ttctaacaaa cctagagttg actttcaacc taaggcatga 3060
gagcgcaaag gctgatctcc tacgatcaca atccatgaat gattcagaag gagacaactg 3120
taaactaaga agaaggaaat tgtcatttga aagtgaagga gagaaagaaa acagtacaaa 3180
tgatcctgaa gactctgcag ataccataag acattatcag agttccaaga gacactttga 3240
agagaaaaaa agcagatcct catctttcat ctcctccatt gatgatgaac aaaagccgct 3300
cttctcagga atagtagact cttctccagg aatagggaaa gcatctgggc tcgattttga 3360
agaaacagtg cccacctcag gaagaatgca catagataaa agaagtcact cttgcaaaga 3420
tatcactgac atgcgaagct gggaacgaga aaatgcacat ccccagcctg aagactccag 3480
tccatctgca cttcagcgag ctgcctgggg tatctctgaa accgaaagcg acctcaccta 3540
cggggaagtg gaacaaagat tagatctgct ccaggagcaa cttaacaggc ttgaatccca 3600
aatgaccact gacatccaga ccatcttaca gttgctgcag aaacaaacca ctgtggtccc 3660
cccagcctac agtatggtaa cagcaggatc agaatatcag agacccatca tccagctgat 3720
gagaaccagt caaccggaag catccatcaa aactgaccga agtttcagcc cttcctcaca 3780
atgtcctgaa tttctagacc ttgaaaaatc taaacttaaa tccaaagaat ccctttcaag 3840
tggggtgcat ctgaacacag cttcagaaga caacttgact tcacttttaa aacaagacag 3900
tgatctctct ttagagcttc acctgcggca aagaaaaact tacgttcatc caattaggca 3960
tccttctttg ccagattcat ccctaagcac tgtaggaatc gtgggtcttc ataggcatgt 4020
ttctgatcct ggtcttccag ggaaataa 4048

<210> 39
<211> 1539
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 7475338CB1

<400> 39
atggagaaca aagaggcggg aaccgctcca cccattccat ccagggaggg gcggctccag 60
ccgacgctgt tgctggcgac actgagcgcg gcctttggct cagccttcca gtacggctac 120
aacctctctg tggtcaacac gccgcacaag gtgttcaagt cattttacaa cgaaacctac 180
tttgagcgac acgcaacatt catggacggg aagctcatgc tgcttctatg gtcttgcacc 240
gtctccatgt ttcctctggg cggcctgttg gggtcattgc tcgtgggcct gctggttgat 300
agctgcggca ggaaggggac cctgctgatc aacaacatct ttgccatcat ccccgccatc 360
ctgatgggag tcagcaaagt ggccaaggct tttgagctga tcgtcttttc ccgagtggtg 420
ctgggagtct gtgcaggtat ctcctacagc gcccttccca tgtacctggg agaactggcc 480
cccaagaacc tgagaggcat ggtgggaaca atgaccgagg ttttcgtcat cgttggagtc 540
ttcctagcac agatcttcag cctccaggcc atcttgggca acccggcagg ttggccggtg 600
cttctggcgc tcacaggggt gcccgccctg ctgcagctgc tgaccctgcc cttcttcccc 660
gaaagccccc gctactccct gattcagaaa ggagatgaag ccacagcgcg acaagctctg 720
aggaggctga gaggccacac ggacatggag gccgagctgg aggacatgcg tgcggaggcc 780
cgggccgagc gcgccgaggg ccacctgtct gtgctgcacc tctgtgccct gcggtccctg 840
cgctggcagc tcctctccat catcgtgctc atggccggcc agcagctgtc gggcatcaat 900
gcgatcaact actatgcgga caccatctac acatctgcgg gcgtggaggc cgctcactcc 960
caatatgtaa cggtgggctc tggcgtcgtc aacatagtga tgaccatcac ctcggctgtc 1020
cttgtggagc ggctgggacg gcggcacctc ctgctggccg gctacggcat ctgcggctct 1080
gcctgcctgg tgctgacggt ggtgctccta ttccagaaca gggtccccga gctgtcctac 1140
ctcggcatca tctgtgtctt tgcctacatc gcgggacatt ccattgggcc cagtcctgtc 1200
ccctcggtgg tgaggaccga gatcttcctg cagtactccc ggcgggcagc tttcatggtg 1260
gacggggcag tgcactggct caccaacttc atcataggct tcctgttccc atccatccag 1320
gaggccatcg gtgcctacag tttcatcatc tttgccggaa tctgcctcct cactgcgatt 1380
tacatctacg tggttattcc ggagaccaag ggcaaaacat ttgtggagat aaaccgcatt 1440
tttgccaaga gaaacagggt gaagcttcca gaggagaaag aagaaaccat tgatgctggg 1500
cctcccacag cctctcctgc caaggaaact tccttttag 1539

<210> 40
<211> 3114
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7476747CB1

<400> 40
ccaagcagtg cctcacttct gccttgtcta gctgtactct ggaaaattaa gaaatttatg 60
agtgtagcac caagtatacc aatgggaagg atgggagtca gaagtcaagt gaactcagcc 120
cgcctctgtg tactttgcac ttttccattt cccttggtac caggcacttt catacttaat 180
ccatagtgga gctgtcacag tgagcaactc tgacaatgac agcttctacc ccagaggcga 240
ccccaaacat ggagctaaag gctccagctg caggaggtct taatgctggc cctgtccccc 300
cagctgccat gtccacgcag agacttcgga atgaagacta ccacgactac agctccacgg 360
acgtgagccc tgaggagagc ccgtcggaag gcctcaacaa cctctcctcc ccgggctcct 420
accagcgctt tggtcaaagc aatagcacaa catggttcca gaccttgatc cacctgttaa 480
aaggcaacat tggcacagga ctcctgggac tccctctggc ggtgaaaaat gcaggcatcg 540
tgatgggtcc catcagcctg ctgatcatag gcatcgtggc cgtgcactgc atgggtatcc 600
tggtgaaatg tgctcaccac ttctgccgca ggctgaataa atcctttgtg gattatggtg 660
atactgtgat gtatggacta gaatccagcc cctgctcctg gctccggaac cacgcacact 720
ggggaagacg tgttgtggac ttcttcctga ttgtcaccca gctgggattc tgctgtgtct 780
attttgtgtt tctggctgac aactttaaac aggtgataga agcggccaat gggaccacca 840
ataactgcca caacaatgag acggtgattc tgacgcctac catggactcg cgactctaca 900
tgctctcctt cctgcccttc ctggtgctgc tggttttcat caggaacctc cgagccctgt 960
ccatcttctc cctgttggcc aacatcacca tgctggtcag cttggtcatg atctaccagt 1020
tcattgttca gaggatccca gaccccagcc acctccccct ggtggccccc tggaagacct 1080
accctctctt ctttggcaca gcgatttttt catttgaagg cattggaatg gttctgcccc 1140
tggaaaacaa aatgaaggat cctcggaagt tcccactcat cctgtacctg ggcatggtca 1200
tcgtcaccat cctctacatc agcctggggt gtctggggta cctgcaattt ggagctaata 1260
tccaaggcag cataaccctc aacctgccca actgctggtt gtaccagtca gttaagctgc 1320
tgtactccat cgggatcttt ttcacctacg cactccagtt ctacgtcccg gctgagatca 1380
tcatcccctt ctttgtgtcc cgagcgcccg agccctgtga gttagtggtg gacctgtttg 1440
tgcgcccagt gctggtctgc ctgacatcac tgtctggcag tgttgacaat ggctggtatg 1500
gcacggaagc cgatggcacc tcctgcggca gtgcaccatt ggtcttcgtc agttcctcct 1560

tcctggctca cccgtggctg agtttcagat gtgagagcca gtgggtgtcc tgtcacagag 1620
atacggtcgt cgtgtggggc ttcgccaggg gcatcttggc catcctcatc ccccgcctgg 1680
acctggtcat ctccctggtg ggctccgtga gcagcagcgc cctggccctc atcatcccac 1740
cgctcctgga ggtcaccacc ttctactcag agggcatgag ccccctcacc atctttaagg 1800
acgccctgat cagcatcctg ggcttcgtgg gctttgtggt ggggacctat gaggctctct 1860
atgagctgat ccagccaagc aatgctccca tcttcatcaa ttccacctgt gccttcatat 1920
agggatcttgg gttcgtctct gcagctgcct acccctgccc catgtgtccc ccgttacctg 1980
tcctcagagc ctcaggtatg gtccaggctc tgaggaaagt cagggttgct gtgtgggaac 2040
ccctctgcct ggcacctgga taccctgggc caggtaacct gagggcagtg gagaggtggg 2100
gtggcagaca cgcagaagtg ctactagtga cagggctgcc atcgctcacc tgtacctatt 2160
tacacccaga actttccagc tccccctcat catgcctcct ccttcctacc tgcctccccc 2220
ctgctggtgc acctcgccca actcattctt actgcacagt tcactttatt taacaatttt 2280
catgtccccc acctcatgtt ttcacctttt ctgggccagg catagattaa gtaactggga 2340
acgccccctc tttataaagc tgggcttctt tctcatctct ctcccaaatg ttgtatctca 2400
gtattcttcc tattcgagtc tccaggggct ggctggacct acctggtcat ttgaaacagg 2460
cccccaagct ggagttttta atctggactc tctggcttgc tgtgacccct aaggcaatgc 2520
ttctcttccc tggattcctt agtgtgggtc acagtactgt gttcttagtt gctttagctc 2580
ttaaaacata cgaagtgttg cctaaactga aaatatttat cttttattta aaatcagatt 2640
tttgttttta gactgtctta gatctggggc tattacgaat cacttcttct tcagtaaact 2700
ttgactcaac ttctcctgct gaaaagaagc tcgctccaga tgtctgcatg ggtcctcggc 2760
actcttggct gaggactcaa aggttttaat caggatcgtc taaaaatgta cctcggtgag 2820
gaggcacaga ttttgcctcc tgttgaccag cctggtttca taccgaaaag acattgaagg 2880
actgcagaaa tgtatgggtg caccgggccg agggaagggt ggctgagtga gaggcgtata 2940
aaatggggct gtgtgcatgc aggcccatgt ttcagcctca gcccacgcca ggtgaaagga 3000
tcagcaatgc tctgttgcca tcgtgctggg acgacaccag ctctattgcc accgatgagt 3060
agctgaggtc agtgtgcaca gagtttgaaa ttaagttaat agactttaca gcag 3114

<210> 41
<211> 2877
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477898CB1

<400> 41
atgccggtcc gcaggggcca cgtcgctccc caaaacactt acctggacac catcatccgc 60
aagttcgagg gccaaagtcg gaagttcctg attgccaatg ctcagatgga gaactgcgcc 120
atcatttact gcaacgacgg cttctgcgaa ctcttcggct actcccgagt ggaggtgatg 180
cagcaaccct gcacctgcga cttcctcaca ggccccaaca caccaagcag cgccgtgtcc 240
cgcctagcgc aggccctgct gggggctgag gagtgcaagg tggacatcct ctactaccgc 300
aaggatgcct ccagcttccg ctgcctggta gatgtggtgc ccgtgaagaa cgaggacggg 360
gctgtcatca tgttcattct caacttcgag gacctggccc agctcctggc caagtgcagc 420
agccgcagct tgtcccagcg cctgttgtcc cagagcttcc tgggctccga gggctctcat 480
ggcaggccag gcggaccagg gccaggcaca ggcaggggca agtacaggac catcagccag 540
atcccacagt tcacgctcaa cttcgtggag ttcaacttgg agaagcaccg ctccagctcc 600
accacggaga ttgagatcat cgcgccccat aaggtggtgg agcggacaca gaacgtcact 660
gagaaggtca cccaggtcct gtccctgggc gcggatgtgc tgccggagta caagctgcag 720
gcgccgcgca tccaccgctg gaccatcctg cactacagcc ccttcaaggc cgtgtgggac 780
tggctcatcc tgctgctggt catctacacg gctgtcttca cgccctactc agccgccttc 840
ctgctcagcg atcaggacga atcacggcgt ggggcctgca gctatacctg cagtcccctc 900
actgtggtgg atctcatcgt ggacatcatg ttcgtcgtgg acatcgtcat caacttccgc 960
accacctatg tcaacaccaa tgatgaggtg gtcagccacc cccgccgcat cgccgtccac 1020
tacttcaagg gctggttcct cattgacatg gtggccgcca tcccttttcga cctcctgatc 1080
ttccgcactg gctccgatga gaccacaacc ctgattgggc tattgaagac agcgcggctg 1140
ctgcggctgg tgcgcgtagc acggaagctg gaccgctact ctgagtatgg ggcggctgtg 1200
ctcttcttgc tcatgtgcac ctttccgctc atagcgcact ggctggcctg catctggtac 1260
gccatcggca atgtggagcg gccctaccta gaacacaaga tcggctggct ggacagcctg 1320
ggtgtgcagc ttggcaagcg ctacaacggc agcgacccag cctcgggccc ctcggtgcag 1380
gacaagtatg tcacagccct ctacttcacc ttcagcagcc tcaccagcgt gggcttcggc 1440
aatgtctcgc ccaacaccaa ctccgagaag gtcttctcca tctgcgtcat gctcatcggc 1500
tccctgatgt acgccagcat cttcgggaac gtgtccgcga tcatccagcg cctgtactcg 1560
ggcaccgcgc gctaccacac gcagatgctg cgtgtcaagg agttcatccg cttccaccag 1620
atccccaacc cactgcgcca gcgcctggag gagtatttcc agcacgcctg gtcctacacc 1680
aatggcattg acatgaacgc ggtgctgaag ggcttccccg agtgcctgca ggctgacatc 1740
tgcctgcacc tgcaccgcgc actgctgcag cactgcccag ctttcagcgg cgccggcaag 1800

ggctgcctgc gcgcgctagc cgtcaagttc aagaccaccc acgcgccgcc tggggacacg 1860
ctggtgcacc tcggcgacgt gctctccacc ctctacttca tctcccgagg ctccatcgag 1920
atcctgcgcg acgacgtggt cgtggccatc ctaggaaaga atgacatctt tggggaaccc 1980
gtcagcctcc atgcccagcc aggcaagtcc agtgcagacg tgcgggctct gacctactgc 2040
gacctgcaca agatccagcg ggcagatctg ctggaggtgc tggacatgta cccggccttt 2100
gcggagagct tctggagtaa gctggaggtc accttcaacc tgcgggacgt aaccgggggt 2160
ctccactcat ccccccgaca ggctcctggc agccaagacc accaaggttt ctttctcagt 2220
gacaaccagt cagatgcagc ccctcccctg agcatctcag atgcattctg gctctggcct 2280
gagctactgc aggaaatgcc cccaaagcac agcccccaaa gccctcagga agacccagat 2340
tgctggcctc tgaagctggg ctccaggcta gagcagctcc aggcccagat gaacaggctg 2400
gagtcccgcg tgtcctcaga cctcagccgc atcttgcagc tcctccagaa gcccatgccc 2460
cagggccacg ccagctacat tctggaagcc actgcctcca atgacctggc cttggttcct 2520
atagcctcgg agacgacgag tccagggccc aggctgcccc agggctttct gcctcctgca 2580
cagaccccaa gctatggaga cttggatgac tgtagtccaa agcacaggaa ctcctccccc 2640
aggatgcctc acctggctgt ggcaatggac aaaactctgg caccatcctc agaacaggaa 2700
cagcctgagg ggctctggcc acccctagcc tcacctctac atcccctgga agtacaagga 2760
ctcatctgtg gtccctgctt ctcctccctc cctgaacacc ttggctctgt tcccaagcag 2820
ctggacttcc agagacatgg ctcagatcct ggatttgcag ggagttgggg ccactga 2877

<210> 42
<211> 2820
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472728CB1

<400> 42
atggggcatc aagggccatt tgaagaagga aatggtggac tgagagtgat agcgacctgg 60
aggaggaagg aggcttggag aagggactgt cttttaggag ccctgcccag tgtttcctgt 120
ggagggtggg gccatcgtgg aagacagacc tatggtaggg cttgtggggt gaaagaaaag 180
ccctttagtc ttttgggtcc tcaaatcaca gtttatgcag tttggcccca gtcagaggga 240
ccccaggaag gcagactcag ggtaaattct gcctgtcttc caccagagag gggactcacc 300
aacgcttgta caaaccatga agaactctct ttggactgtt tgctttttga gaatgttaac 360
accttgactc tggatttctg cctatgggaa aaaaccacaa tagtgccagg ggtgcttcca 420
tatgcaggat taactctgca gtcaaagttt ctgttgggca gagcattgtt agcaggggtc 480
catgtgatca cactgacacc tgagagagtg acacaccatg tacatggctg gtatatggag 540
gatggattta agggggacag gactgaaggc tgtcgcagtg attcagtggc cgttcccgca 600
gcagcaccgg tgtgccagcc caagagcgcc actaacgggc aacccccggc tccggctccg 660
actccaactc cgcgcctgtc catttcctcc cgagccacag tggtagccag gatggaaggc 720
acctcccaag ggggcttgca gaccgtcatg aagtggaaga cggtggttgc catctttgtg 780
gttgtggtgg tctaccttgt cactggcggt cttgtcttcc gggcattgga gcagcccttt 840
gagagcagcc agaagaatac catcgccttg gagaaggcgg aattcctgcg ggatcatgtc 900
tgtgtgagcc cccaggagct ggagacgttg atccagcatg ctcttgatgc tgacaatgcg 960
ggagtcagtc caataggaaa ctcttccaac aacagcagcc actgggacct cggcagtgcc 1020
tttttctttg ctggaactgt cattacgacc atgtatggga atattgctcc gagcactgaa 1080
ggaggcaaaa tcttttgtat tttatatgcc atctttggaa ttccactctt tggtttctta 1140
ttggctggaa ttggagacca acttggaacc atctttggga aaagcattgc aagagtggag 1200
aaggtctttc gaaaaaagca agtgagtcag accaagatcc gggtcatctc aaccatcctg 1260
ttcatcttgg ccggctgcat tgtgtttgtg acgatccctg ctgtcatctt taagtacatc 1320
gagggctgga cggccttgga gtccatttac tttgtggtgg tcactctgac cacggtgggc 1380
tttggtgatt ttgtggcagt ggttgttttc aggggaaacg ctggcatcaa ttatcgggag 1440
tggtataagc ccctagtgtg gttttggatc cttgttggcc ttgcctactt tgcagctgtc 1500
ctcagtatga tcggagattg gctacgggtt ctgtccaaaa agacaaaaga agaggtgggt 1560
gaaatcaagg cccatgcggc agagtggaag gccaatgtca cggctgagtt ccgggagaca 1620
cggcgaaggc tcagcgtgga gatccacgat aagctgcagc gggcagccac catccgcagc 1680
atggagcgcc ggcggctggg cctggaccag cgggcccact cactggacat gctgtccccc 1740
gagaagcgct ctgtctttgc tgccctggac accggccgct tcaaggcctc atcccaggag 1800
agcatcaaca accggcccaa caacctgcgc ctgaaggggc cggagcagct gaacaagcat 1860
gggcagggtg cgtccgagga caacatcatc aacaagttcg ggtccacctc cagactcacc 1920
aagaggaaaa acaaggacct caaaaagacc ttgcccgagg acgttcagaa aatctacaag 1980
accttccgga attactccct ggacgaggag aagaaagagg aggagacgga aaagatgtgt 2040
aactcagaca actccagcac agccatgctg acggactgta tccagcagca cgctgagttg 2100
gagaacggaa tgatacccac ggacaccaaa gaccgggagc cggagaacaa ctcattactt 2160
gaagacagaa actaaatgtg aaggacattg gtcttggact gagcgttgtg tgtgtgtgtg 2220
tgtgtgtgtt tttaatattc acactgagac atgtgcctta aacagacttt ttagtccaaa 2280

attacatagc attgaagaat atatttcact gtgccataaa caactgaaag cttgctctgc 2340
caaaaggaat cagagaacaa gaacttcatt tcagatagca aacgcaggac acaccaagag 2400
tgtccgtgca cgtagccggt tctggacgta catgttaagg gcatttcagt ggcagtgctg 2460
tacccctggg cagtgctacc tgggcacaca cgtagacaag ggcagctatt ccttagacca 2520
gcctcctgaa agaaacaggt gtgtcttttt agtggagtcg tagtaatatg tgcacacaca 2580
gaaggggacc tgattgggtg ggagctggtt atgtgtaact agcgttggag ttgacatttt 2640
ggcatgtgct ctgagcttga attttgatac caaccattca gtgcatcata cctagtcttt 2700
ctatgctcca aatgaatgtc tgtggggacc tgagagcacc tggaatttgt tggaagcaga 2760
tcagagcaca cgtacgaaaa ggtgcaattc cttttctcat gacaaaaggg aaaaaaataa 2820

<210> 43
<211> 1440
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474322CB1

<220>
<221> unsure
<222> 757
<223> a, t, c, g, or other

<400> 43
atgtacaatg agattctgat gctgggggcc aaactgcacc cgacgctgaa gctggaggag 60
ctcaccaaca agaagggaat gacgccgctg gctctggcag ctgggaccgg gaagatcggg 120
aatcgccacg acatgctctt ggtggagccg ctgaaccgac tcctgcagga caagtgggac 180
agattcgtca agcgcatctt ctacttcaac ttcctggtct actgcctgta catgatcatc 240
ttcaccatgg ctgcctacta caggcccgtg gatggcttgc ctccctttaa gatggaaaaa 300
actggagact atttccgagt tactggagag atcctgtctg tgttaggagg agtctacttc 360
tttttccgag ggattcagta tttcctgcag aggcggccgt cgatgaagac cctgtttgtg 420
gacagctaca gtgagatgct tttgtttctg cagtcactgt tcatgctggc caccgtggtg 480
ctgtacttca gccacctcaa ggagtatgtg gcttccatgg tattctccct ggccttgggc 540
tggaccaaca tgctctacta cacccgcggt ttccagcaga tgggcatcta tgccgtcatg 600
atagagaaga tgatcctgag agacctgtgc cgtttcatgt ttgtctacat cgtcttcttg 660
ttcgggtttt ccacagcggt ggtgacgctg attgaagacg ggaagaatga ctccctgccg 720
tctgagtcca cgtcgcacag gtggcggggg cctgctngca ggcccaatag ctcctacaac 780
agcctgtact ccacctgcct ggagctgttc aagttcacca tcggcatggg cgacctggag 840
ttcactgaga actatgactt caaggctgtc ttcatcatcc tgctgctggc ctatgtaatt 900
ctcacctaca tagttctcct cctcaacatg ctcattgctc tgatgggcga gactgtggag 960
aacgtctcca aggagagcga acgcatctgg cgcctgcaga gagccatcac catcctggac 1020
acggagaaga gcttccttaa gtgcatgagg aaggccttcc gctcaggcaa gctgctgcag 1080
gtggggtaca cacctgatgg caaggacgac taccggtggt gcttcgtgga cgaggtgaac 1140
tggaccacct ggaacaccaa cgtgggcatc atcaacgaag acccgggcaa ctgtgagggc 1200
gtcaagcgca ccctgagctt ctccctgcgg tcaagcagag tttcaggcag acactggaag 1260
aactttgccc tggtccccct tttaagagag gcaagtgctc gagataggca gtctgctcag 1320
cccgaggaag tttatctgcg acagttttca gggtctctga agccagagga cgctgaggtt 1380
ttcaagagtc ctgccgcttc cggggagaag tgaggacgtc acgcagacag cactgtcaac 1440

<210> 44
<211> 2394
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5455621CB1

<400> 44
atttcagaac acatctgaat tccttctctg tggcatatgc tttaggagag gagcagacag 60
ctcttagcta gggtcagatt tcaaattctc atctcttggt gccaatacca ccaccagatt 120
cttctttgaa gtcaactttt gagatcttca ctaagtacac gttggtgtct gaagattcac 180
acgagtgcct ctggtaatca ttttcttcag ggaatcacag tctctcctct cagcaaagca 240
tccactgtac tgaactttgc ttttggaaac atcttcttcc tgagacctcg ttgaaagaaa 300
ctctctggtg tcatactttc caatatggag gtgaagaact ttgcagtttg ggattatgtt 360
gtatttgcag ccctcttttt catttcctct ggaattgggg tgttctttgc cattaaggag 420

agaaaaaagg caacttcccg agagttcctg gttgggggaa ggcaaatgag ctttggccct 480
gtcggcttgt ctctgacagc cagcttcatg tcagctgtca cggtcctggg gaccccttct 540
gaagtctacc gctttggggc atccttccta gtcttcttca ttgcttacct atttgtcatc 600
ctcttaacat cagagctctt tctccctgtg ttctacagat ctggtatcac cagcacttat 660
gagtacttac aactacgatt caacaaacca gttcgctatg ctgccacagt catctacatt 720
gtacagacga ttctctacac aggagtggtg gtgtatgctc ctgccctggc actcaatcaa 780
gtgactgggt ttgatctctg gggctctgtg tttgcaacag gaattgtttg cacattctac 840
tgtaccctgg gaggattaaa agcagtggtg tggacagatg catttcagat ggttgtcatg 900
attgtgggct tcttaacggt tctcattcaa ggatcaactc atgctggggg attccacaat 960
gtattagagc aatcaacaaa tggatctcga ctacatatat ttgactttga tgtagatcct 1020
ctcaggcgac acacttttg gactatcaca gtgggaggaa cttttacttg gctcggaatc 1080
tatggggtca atcaatcaac tattcagcga tgcatctctt gcaaaacaga aaagcatgct 1140
aagcttgcct tgtattttaa cttgctgggt ctctggatca ttctggtgtg tgctgtcttc 1200
tctggcttaa tcatgtactc tcactttaaa gactgtgacc cttggacttc tggcatcatc 1260
tcagcaccag accagctgat gccgtacttt gtcatggaga tatttgccac aatgccagga 1320
ctgccaggac tttttgtggc ttgtgccttc agtggaactc tgagcaccgt ggcttccagc 1380
atcaatgcct tggcaacagt gacctttgag gattttgtca agagctgttt tcctcatctc 1440
tccgacaagc tgagcacctg gatcagtaaa ggcttatgtc tcttatttgg cgtgatgtgt 1500
acctctatgg ctgtggctgc atctgtcatg ggaggtgttg tgcaggcttc cctcagcatt 1560
cacggcatgt gtggaggacc aatgctgggc ttattctccc tgggaatcgt gttccctttt 1620
gtgaactgga agggtgcact aggaggtctt cttactggaa tcaccttgtc attttgggtg 1680
gccattgggg ccttcattta ccctgcacca gcctctaaga catggccttt gcctctatca 1740
acagaccaat gtatcaaatc aaatgtgaca gcaacagggc ctccagtact atccagcaga 1800
cctggaatag ctgatacctg gtactcgatc tcctaccttt actacagtgc actgggctgc 1860
ttaggatgca ttgttgctgg agtaatcatc agcctcataa caggtcgcca aagaggtgag 1920
gatattcaac cactgttaat tagaccagtt tgtaatttat tttgcttttg gtctaagaag 1980
tacaaaacac tatgctggtg tggagttcag catgacagtg ggacagagca ggaaaacctt 2040
gagaatggca gtgcccggaa acaggggget gaatctgtct tacagaacgg actcagaaga 2100
gaaagcctgg tacatgttcc aggctatgat cctaaggaca aaagctacaa caatatggca 2160
tttgagacta cccatttcta aggcaatacc tgtatgaatg cacacacaca cgtgcaatac 2220
acacacacac acacaaactc cacatacttc ttgcctactt gttagtagat atgtatagtt 2280
gccattgcta gaagacaggg atgtctggtg cctatttcta cttatttata actacatgca 2340
aaatgactgt ctctcgggat attctttgaa agactccaac tttcacagag aaaa 2394

<210> 45
<211> 2890
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477248CB1

<400> 45
gaatactaag ccagggcaga atgcttgtga agtagcaact aaagtggcag tgtttcttct 60
gaaattctca ggcagtcaga ctgtcttagg caaatcttga taaaatagcc cttatccagg 120
tttttatcta aggaatccca agaagactgg ggaatggaga gacagtcaag ggttatgtca 180
gaaaaggatg agtatcagtt tcaacatcag ggagcggtgg agctgcttgt cttcaatttt 240
ttgctcatcc ttaccatttt gacaatctgg ttatttaaaa atcatcgatt ccgcttcttg 300
catgaaactg gaggagcaat ggtgtatggc cttataatgg gactaatttt acgatatgct 360
acagcaccaa ctgatattga aagtggaact gtctatgact gtgtaaaact aactttcagt 420
ccatcaactc tgctggttaa tatcactgac caagtttatg aatataaata caaaagagaa 480
ataagtcagc acaacatcaa tcctcatcaa ggaaatgcta tacttgaaaa gatgacattt 540
gatccagaaa tcttcttcaa tgttttactg ccaccaatta tatttcatgc aggatatagt 600
ctaaagaaga gacacttttt tcaaaactta ggatctattt taacgtatgc cttcttggga 660
actgccatct cctgcatcgt catagggtta attatgtatg gttttgtgaa ggctatgata 720
catgctggcc agctgaaaaa tggagacttt catttcactg actgtttatt ttttggttca 780
ctgatgtctg ctacagatcc agtgacagtg ctggccattt tccatgaact gcacgtcgac 840
cctgacctgt acacactctt gtttggagag agtgtgttga atgatgcagt ggccatagtc 900
cttacatatt ctatatccat ttacagtccc aaggagaatc caaatgcatt tgatgccgca 960
gcattcttcc agtctgtggg gaatttcctg ggaatcttcg ctggctcatt tgcaatgggg 1020
tctgcgtatg ccatcatcac agcactgttg accaaattta ccaagctgtg tgagttcccg 1080
atgctggaaa ccggcctgtt tttcctgctt tcttggagtg ccttcctgtc tgccgaggct 1140
gccggcctaa cagggatagt tgctgttctc ttctgtggag tcacacaagc acattatacc 1200
tacaacaatc tgtcttcgga ttccaaaata agaactaaac agttgtttga atttatgaac 1260
tttttggcgg agaacgtcat cttctgttac atgggcctgg cactgttcac gttccagaat 1320
catatcttta atgctctttt tatacttgga gcctttctag caatttttgt tgccagagcc 1380

tgcaacatat atcccctctc cttcctcctg aatctaggcc gaaaacagaa gatcccctgg 1440
aactttcagc acatgatgat gttttcaggt ttgcgaggag cgatcgcatt tgccttagct 1500
attcggaaca cagaatctca gctcaaacaa atgatgttta ccactacgct gctcctcgtg 1560
ttcttcactg tctgggtatt tggaggagga acaaccccca tgttgacttg gcttcagatc 1620
agagttggcg tggacctgga tgaaaatctg aaggaggacc cctcctcaca acaccaggaa 1680
gcaaataact tggataaaaa catgacgaaa gcagagagtg ctcggctctt cagaatgtgg 1740
tatagctttg accacaagta tctgaaacca attttaaccc actctggtcc tccgctgact 1800
acaacattac ctgaatggtg tggtccgatt tccaggctgc ttaccagtcc tcaagcctat 1860
ggggaacagc taaaagagga tgatgtggaa tgcattgtaa accaggatga actagccata 1920
aattaccagg agcaagcctc ctcaccctgc agtcctcctg caaggctagg tctggaccag 1980
aaagcttcac cccagacgcc aggcaaggaa aacatttatg agggagacct cggccctggg 2040
aggctatgaa ctcaagcttg agcaaacttt gggtcaatcc cagttgaatt aattggcatg 2100
aagagtacag atgtaatcac aagtaatgca agactcactg aggaatacaa gccaagctga 2160
tgaggcagta caggggagag gctggaaaac atattaagag cataaattgg agagaatcaa 2220
agccttgtca catggatcct ctggtgcctg aagaaatgag attttattat ccctctctat 2280
tatgcaaatg aatttagttt tttgacagca gccattctga ttactggatt ggctggggtg 2340
gggatggagg tatcaggagt ctagctgctg gaggatggga cagctgtgct gggtcttcag 2400
ggcatttctg ctgcgaatgc ggctctccag gcccttcact tctattctgg attttattcc 2460
ctccattaag gagagtttaa aaataaaaga aagcttctga gagtaaacat tttgctccta 2520
agctgaaggg aatgcccagc tatttagtaa gtgataagtt tcttattttg aggacttgac 2580
tcccatttgc tctcagtgac cccagggcag agcccagaga agtgttccgt acccactgct 2640
gatggtttcc cagagcccac actgagttga agaacctatt gttcttcttg gcatccttct 2700
tatgctactt ctcccatcgc tcaaaggggt tgcctatggc tgggtgtgcc ctgccctaaa 2760
tgcagcacca ctttcaagca gcttctagct atagctttcc accaggtatt tttaatccca 2820
tttcacctcc tcccccagca attcaccagt caggagtgat ttttactgta aagatggttg 2880
cttagtaaaa 2890

<210> 46
<211> 3926
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2944004CB1

<400> 46
ctggaccttt aatccactgt aggtatggac agggaagaaa ggaagaccat caatcagggt 60
caagaagatg aaatggagat ttatggttac aatttgagtc gctggaagct tgccatagtt 120
tctttaggag tgatttgctc tggtggggtt tctcctcctc ctctctattg gatgcctgag 180
tggcgggtga aagcgacctg tgtcagagct gcaattaaag actgtgaagt agtgctgctg 240
aggactactg atgaattcaa aatgtggttt tgtgcaaaaa ttcgcgttct ttctttggaa 300
acttacccag tttcaagtcc aaaatctatg tctaataagc tttcaaatgg ccatgcagtt 360
tgtttaattg agaatcccac tgaagaaaat aggcacagga tcagtaaata ttcacagact 420
gaatcacaac agattcgtta tttcacccac catagtgtaa aatatttctg gaatgatacc 480
attcacaatt ttgatttctt aaagggactg gatgaaggtg tttcttgtac gtcaatttat 540
gaaaagcata gtgcaggact gacaaagggg atgcatgcct acagaaaact gctttatgga 600
gtaaatgaaa ttgctgtaaa agtgccttct gtttttaagc ttctaattaa agaggttctc 660
aacccatttt acattttcca gctgttcagt gttatactgt ggagcactga tgaatactat 720
tactatgctc tagctattgt ggttatgtcc atagtatcaa tcgtaagctc actatattcc 780
attagaaagc aatatgttat gttgcatgac atggtggcaa ctcatagtac cgtaagagtt 840
tcagtttgta gagtaaatga agaaatagaa gaaatctttt ctaccgacct tgtgccagga 900
gatgtcatgg tcattccatt aaatgggaca ataatgcctt gtgatgctgt gcttattaat 960
ggtacctgca ttgtaaacga aagcatgtta acaggagaaa gtgttccagt gacaaagact 1020
aatttgccaa atccttcagt ggatgtgaaa ggaataggag atgaattata taatccagaa 1080
acacataaac gacatacttt gttttgtggg acaactgtta ttcagactcg tttctacact 1140
ggagaactcg tcaaagccat agttgttaga acaggattta gtacttccaa aggacagctt 1200
gttcgttcca tattgtatcc caaaccaact gattttaaac tctacagaga tgcctacttg 1260
tttctactat gtcttgtggc agttgctggc attgggttta tctacactat tattaatagc 1320
attttaaatg aggtacaagt tggggtcata attatcgagt ctcttgatat tatcacaatt 1380
actgtgcccc ctgcacttcc tgctgcaatg actgctggta ttgtgtatgc tcagagaaga 1440
ctgaaaaaaa tcggtatttt ctgtatcagt cctcaaagaa taaatatttg tggacagctc 1500
aatcttgttt gctttgacaa gactggaact ctaactgaag atggtttaga tctttggggg 1560
attcaacgag tggaaaatgc acgatttctt tcaccagaag aaaatgtgtg caatgagatg 1620
ttggtaaaat cccagtttgt tgcttgtatg gctacttgtc attcacttac aaaaattgaa 1680
ggagtgctct ctggtgatcc acttgatctg aaaatgtttg aggctattgg atggattctg 1740
gaagaagcaa ctgaagaaga aacagcactt cataatcgaa ttatgcccac agtggttcgt 1800

cctcccaaac aactgcttcc tgaatctacc cctgcaggaa accaagaaat ggagctgttt 1860
gaacttccag ctacttatga gataggaatt gttcgccagt tcccattttc ttctgctttg 1920
caacgtatga gtgtggttgc cagggtgctg ggggatagga aaatggacgc ctacatgaaa 1980
ggagcgcccg aggccattgc cggtctctgt aaacctgaaa cagttcctgt cgattttcaa 2040
aacgttttgg aagacttcac taaacagggc ttccgtgtga ttgctcttgc acacagaaaa 2100
ttggagtcaa aactgacatg gcataaagta cagaatatta gcagagatgc aattgagaac 2160
aacatggatt ttatgggatt aattataatg cagaacaaat taaagcaaaa aacccctgca 2220
gtacttgaag atttgcataa agccaacatt cgcaccgtca tggtcacagg tgacagtatg 2280
ttgactgctg tctctgtggc cagagattgt ggaatgattc tacctcagga taaagtgatt 2340
attgctgaag cattacctcc aaaggatggg aaagttgcca aaataaattg gcattatgca 2400
gactccctca cgcagtgcag tcatccatca gcaattgacc cagaggctat tccggttaaa 2460
ttggtccatg atagcttaga ggatcttcaa atgactcgtt atcattttgc aatgaatgga 2520
aaatcattct cagtgatact ggagcatttt caagaccttg ttcctaagtt gatgttgcat 2580
ggcaccgtgt ttgcccgtat ggcacctgat cagaagacac agttgataga agcattgcaa 2640
aatgttgatt attttgttgg gatgtgtggt gatggcgcaa atgattgtgg tgctttgaag 2700
agggcacacg gaggcatttc cttatcggag ctcgaagctt cagtggcatc tccctttacc 2760
tctaagactc ctagtatttc ctgtgtgcca aaccttatca gggaaggccg tgctgcttta 2820
ataacttcct tctgtgtgtt taaattcatg gcattgtaca gcattatcca gtacttcagt 2880
gttactctgc tgtattctat cttaagtaac ctaggagact tccagtttct cttcattgat 2940
ctggcaatca ttttggtagt ggtatttaca atgagtttaa atcctgcctg gaaagaactt 3000
gtggcacaaa gaccaccttc gggtcttata tctggggccc ttctcttctc cgttttgtct 3060
cagattatca tctgcattgg atttcaatct ttgggttttt tttgggtcaa acagcaacct 3120
tggtatgaag tgtggcatcc aaaatcagat gcttgtaata caacaggaag cgggttttgg 3180
aattcttcac acgtagacaa tgaaaccgaa cttgatgaac ataatatacA aaattatgaa 3240
aataccacag tgtttttat ttccagtttt cagtacctca tagtggcaat tgccttttca 3300
aaaggaaaac ccttcaggca accttgctac aaaaattatt tttttgtttt ttctgtgatt 3360
tttttatata tttttatatt attcattatg ttgtatccag ttgcctctgt tgaccaggtt 3420
cttcagatag tgtgtgtacc atatcagtgg cgtgtaacta tgctcatcat tgttcttgtc 3480
aatgcctttg tgtctatcac agtggagaac ttcttccttg acatggtcct ttggaaagtt 3540
gtgttcaacc gagacaaaca aggagagtat cggttcagca ccacacagcc accgcaggag 3600
tcagtggatc ggtggggaaa atgctgctta ccctgggccc tgggctgtag aaagaagaca 3660
ccaaaggcaa agtacatgta tctggcgcag gagctcttgg ttgatccaga atggccacca 3720
aaacctcaga caaccacaga agctaaagct ttagttaagg agaatggatc atgtcaaatc 3780
atcaccataa catagcagtg aatcagtctc agtggtattg ctgatagcag tattcaggaa 3840
tatgtgattt taggagtttc tgatcctgtg tgtcagaatg gcactagttc agtttatgtc 3900
ccttctgata tagtagctta tttgac 3926

<210> 47
<211> 2135
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3046849CB1

<400> 47
cgctcaggcc cctctttcga atgctccacg ccctcctgcg atctagaacg attcagggca 60
ggatcctgct cctgaccatc tgcgctgccg gcattggtgg gacttttcag tttggctata 120
acctctctat catcaatgcc ccgaccttgc acattcagga attcaccaat gagacatggc 180
aggcgcgtac tggagagcca ctgcccgatc acctagtcct gcttatgtgg tccctcatcg 240
tgtctctgta tcccctggga ggcctctttg gagcactgct tgcaggtccc ttggccatca 300
cgctgggaag gaagaagtcc ctcctggtga ataacatctt tgtggtgtca gcagcaatcc 360
tgtttggatt cagccgcaaa gcaggctcct ttgagatgat catgctggga agactgctcg 420
tgggagtcaa tgcaggtgtg agcatgaaca tccagcccat gtacctgggg gagagcgccc 480
ctaaggagct ccgaggagct gtggccatga gctcagccat ctttacggct ctggggatcg 540
tgatgggaca ggtggtcgga ctcagggagc tcctaggtgg ccctcaggcc tggcccctgc 600
tgctggccag ctgcctggtg cccggggcgc tccagctcgc ctccctgcct ctgctccctg 660
aaagcccgcg ctacctcctc attgactgtg gagacaccga ggcctgcctg gcagcactac 720
ggcagctacg gggctccggg gacttggcag gggagctgga ggagctggag gaggagcgcg 780
ctgcctgcca gggctgccgt gcccggcgcc catgggagct gttccagcat cgggccctga 840
ggagacaggt gacaagcctc gtggttctgg gcagtgccat ggagctctgc gggaatgact 900
cggtgtacgc ctacgcctcc tccgtgttcc ggaaggcagg agtgccggaa gcgaagatcc 960
agtacgcgat catcgggact gggagctgcg agctgctcac ggcggttgtt agttgtgtgg 1020
taatcgagag ggtgggtcgg cgcgtgctgc tcatcggtgg gtacagcctg atgacctgct 1080
gggggagcat cttcactgtg gccctgtgcc tgcagagctc cttcccctgg acactctacc 1140
tggccatggc ctgcatcttt gccttcatcc tcagctttgg cattggccct gccggagtga 1200

cggggatcct ggccacagag ctgtttgacc agatggccag gcctgctgcc tgcatggtct 1260
gcggggcgct catgtggatc atgctcatcc tggtcggcct gggatttccc tttatcatgg 1320
aggccttgtc ccacttcctc tatgtccctt tccttggtgt ctgtgtctgt ggggccatct 1380
acactggcct gttccttcct gagaccaaag gcaagacctt ccaagagatc tccaaggaat 1440
tacacagact caacttcccc aggcgggccc agggccccac gtggaggagc ctggaggtta 1500
tccagtcaac agaactctag tcccaaaggg gtggccagag ccaaagccag ctactgtcct 1560
gtcctctgct tcctgccagg gccctggtcc tcactccctc ctgcattcct catttaagga 1620
gtgtttattg agcacccttt gtgtgcagac atggctccag gtgcttagca atcaatggtg 1680
agcgtggtat tccaggctaa aggtaattaa ctgacagaaa atcagtaaca acataattac 1740
aggctggttg tggcagctca tgactgtaat cccagcactt tgggaggcca aggtgggagg 1800
atcaattgag gccagagttt gaaaccagcc taggtaacat agtgagaccc cctatctcta 1860
caaaaaattt taaacattag ctgggcatgg tggtatgtgc taacagctct agctactcag 1920
gaggctgagg cagcaggatc acttgagtcc caagagttca aggtagcagt aagctaacaa 1980
ttcacaccac tgcatgccca gactggggtg acagagggag acttcatctc tttaaaaaca 2040
taataataat aattacggac tccggaaatg cgttgacaac gaaacatacc ggtggccccg 2100
tgaggtggtg atcccgtatc ccagccttgg gaagc 2135

<210> 48
<211> 2637
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4538363CB1

<400> 48
atgggctgga gatgccactg tccgcttggt ttaatgatca atgagctccc tgccaggaaa 60
ccctttctga cctggtttgc ccctcagtcc ctcgggctca tacctagtgc ctgcggcagg 120
acagccatgg ccgccaactc caccagcgac ctccacactc ccgggacgca gctgagcgtg 180
gctgacatca tcgtcatcac tgtgtatttt gctctgaacg tggccgtggg catatggtcc 240
tcttgtcggg ccagtaggaa cacggtgaat ggctacttcc tggcaggccg ggacatgacg 300
tggtggccga ttggagcctc cctcttcgcc agcagcgagg gctctggcct cttcattgga 360
ctggcgggct caggcgcggc aggaggtctg gccgtggcag gcttcgagtg gaatgccacg 420
tacgtgctgc tggcactggc atgggtgttc gtgcccatct acatctcctc agagatcgtc 480
accttacctg agtacattca gaagcgctac gggggccagc ggatccgcat gtacctgtct 540
gtcctgtccc tgctactgtc tgtcttcacc aagatatcgc tggacctgta cgcgggggct 600
ctgtttgtgc acatctgcct gggctggaac ttctacctct ccaccatcct cacgctcggc 660
atcacagccc tgtacaccat cgcagggggc ctggctgctg taatctacac ggacgccctg 720
cagacgctca tcatggtggt gggggctgtc atcctgacaa tcaaagcttt tgaccagatc 780
ggtggttacg ggcagctgga ggcagcctac gcccaggcca ttccctccag gaccattgcc 840
aacaccacct gccacctgcc acgtacagac gccatgcaca tgtttcgaga cccccacaca 900
ggggacctgc cgtggaccgg gatgaccttt ggcctgacca tcatggccac ctggtactgg 960
tgcaccgacc aggtcatcgt gcagcgatca ctgtcagccc gggacctgaa ccatgccaag 1020
gcgggctcca tcctggccag ctacctcaag atgctcccca tgggcctgat catcatgccg 1080
ggcatgatca gccgcgcatt gttcccagat gatgtgggct gcgtggtgcc gtccgagtgc 1140
ctgcgggcct gcggggccga ggtcggctgc tccaacatcg cctaccccaa gctggtcatg 1200
gaactgatgc ccatcggtct gcgggggctg atgatcgcag tgatgctggc ggcgctcatg 1260
tcgtcgctga cctccatctt caacagcagc agcaccctct tcactatgga catctggagg 1320
cggctgcgtc cccgctccgg cgagcgggag ctcctgctgg tgggacggct ggtcatagtg 1380
gcactcatcg gcgtgagtgt ggcctggatc cccgtcctgc aggactccaa cagcgggcaa 1440
ctcttcatct acatgcagtc agtgaccagc tccctggccc caccagtgac tgcagtcttt 1500
gtcctgggcg tcttctggcg acgtgccaac gagcaggggg ccttctgggg cctgatagca 1560
gggctggtgg tgggggccac gaggctggtc ctggaattcc tgaacccagc cccaccgtgc 1620
ggagagccag acacgcggcc agccgtcctg gggagcatcc actacctgca cttcgctgtc 1680
gccctctttg cactcagtgg tgctgttgtg gtggctggaa gcctgctgac cccaccccca 1740
cagagtgtcc agattgagaa ccttacctgg tggaccctgg ctcaggatgt gcccttggga 1800
actaaagcag gtgatggcca aacaccccag aaacacgcct tctgggcccg tgtctgtggc 1860
ttcaatgcca tcctcctcat gtgtgtcaac atattctttt atgcctactt cgcctgaaca 1920
ctgccatcct ggacagaaag gcaggagctc tgagtcctca ggtccaccca tttccctcat 1980
ggggatcccg aggccccaag aggggcagat tcccctcaca gctgcacagc agctcggtgc 2040
ccaagaactg gccaagccag caaagcggga gcctgaaaac attaggggcg aaactgggac 2100
gaaacataag tgtgactttt tccaaacaac agcacccaaa gcaagtcaag catttggaac 2160
gcgacaaact tagattttcc tgaccgggcc caccacaccc caacctcctc acctcccaaa 2220
ctaccaacac agctcatcac catactcaca ccacccacag cggcccgccc ccactccaat 2280
cagaaaggca ccccccccact ctcaagacgc gacggcgcaa tcgactgcaa ctccataacg 2340
atgccaaaac gacacaagcc aggacacggc actgtataca gcacgagggt gatctgcaac 2400

gttgtggccg aatgcagaaa atacactggg tgctggcgta aggaagatcc gcgagtaaac 2460
aacggtcttg taaacttact gcatccacca aggtacactt ccagaacgag accagacaac 2520
tacactccac acaacctgca gccacaccct atttctgcta tcataaagag cccccgcacc 2580
acataataat gccggcagac tcagtgcgcg aaaccccttgt gctggacttc accacgg 2637

<210> 49
<211> 3783
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 6427460CB1

<400> 49
gcactagtac cccggagccc atgggcgcgc cgagccgggc gcgggggcgc tgaacggcgg 60
agcgggagcg gccggaggag ccatggactg cagcctcgtg cggacgctcg tgcacagata 120
ctgtgcagga gaagagaatt gggtggacag caggaccatc tacgtgggac acagggagcc 180
acctccgggc gcagaggcct acatcccaca gagataccca gacaacagga tcgtctcgtc 240
caagtacaca ttttggaact ttataaccaa gaatttattt gaacaattca gaagagtagc 300
caacttttat ttccttatca tatttctggt gcagttgatt attgatacac ccacaagtcc 360
agtgacaagc ggacttccac tcttctttgt cattactgtg acggctatca aacagggtta 420
tgaagactgg cttcgacata aagcagacaa tgccatgaac cagtgtcctg ttcatttcat 480
tcagcacggc aagctcgttc ggaaacaaag tcgaaagctg cgagttgggg acattgtcat 540
ggttaaggag gacgagacct ttccctgcga cttgatcttc ctttccagca accggggaga 600
tgggacgtgc cacgtcacca ccgccagctt ggatggagaa tccagccata aaacgcatta 660
cgcggtccag gacaccaaag gcttccacac agaggaggat atcggcggac ttcacgccac 720
catcgagtgt gagcagcccc agcccgacct ctacaagttc gtgggtcgca tcaacgttta 780
cagtgacctg aatgaccccg tggtgaggcc cttaggatcg gaaaacctgc tgcttagagg 840
agctacactg aagaacactg agaaaatctt tggtgtggct atttacacgg gaatggaaac 900
caagatggca ttaaattatc aatcaaaatc tcagaagcga tctgccgtgg aaaaatcgat 960
gaatgcgttc ctcattgtgt atctctgcat tctgatcagc aaagccctga taaacactgt 1020
gctgaaatac gtgtggcaga gtgagccctt tcgggatgag ccgtggtata atcagaaaac 1080
ggagtcggaa aggcagagga atctgttcct caaggcattc acggacttcc tggccttcat 1140
ggtcctcttt aactacatca tccctgtgtc catgtacgtc acggtcgaga tgcagaagtt 1200
cctcggctct tacttcatca cctgggacga agacatgttt gacgaggaga ctggcgaggg 1260
gcctctggtg aacacgtcgg acctcaatga agagctggga caggtggagt acatcttcac 1320
agacaagacc ggcaccctca cggaaaacaa catggagttc aaggagtgct gcatcgaagg 1380
ccatgtctac gtgccccacg tcatctgcaa cgggcaggtc ctcccagagt cgtcaggaat 1440
cgacatgatt gactcgtccc ccagcgtcaa cgggagggag cgcgaggagc tgtttttccg 1500
ggccctctgt ctctgccaca ccgtccaggt gaaagacgat gacagcgtag acggccccag 1560
gaaatcgccg gacgggggga aatcctgtgt gtacatctca tcctcgcccg acgaggtggc 1620
gctggtcgaa ggtgtccaga gacttggctt tacctaccta aggctgaagg acaattacat 1680
ggagatatta aacagggaga accacatcga aaggtttgaa ttgctggaaa ttttgagttt 1740
tgactcagtc agaaggagaa tgagtgtaat tgtaaaatct gctacaggag aaatttatct 1800
gttttgcaaa ggagcagatt cttcgatatt cccccgagtg atagaaggca aagttgacca 1860
gatccgagcc agagtggagc gtaacgcagt ggaggggctc cgaactttgt gtgttgctta 1920
taaaaggctg atccaagaag aatatgaagg catttgtaag ctgctgcagg ctgccaaagt 1980
ggcccttcaa gatcgagaga aaaagttagc agaagcctat gagcaaatag agaaagatct 2040
tactctgctt ggtgctacag ctgttgagga ccggctgcag gagaaagctg cagacaccat 2100
cgaggccctg cagaaggccg ggatcaaagt ctgggttctc acgggagaca agatggagac 2160
ggccgcggcc acgtgctacg cctgcaagct cttccgcagg aacacgcagc tgctggagct 2220
gaccaccaag aggatcgagg agcagagcct gcacgacgtc ctgttcgagc tgagcaagac 2280
ggtcctgcgc cacagcggga gcctgaccag agacaacctc tccggacttt cagcagatat 2340
gcaggactac ggtttaatta tcgacggagc tgcactgtct ctgataatga agcctcgaga 2400
agacgggagt tccggcaact acagggagct cttcctggaa atctgccgga gctgcagcgc 2460
ggtgctctgc tgccgcatgg cgcccttgca gaaggctcag attgttaaat taatcaaatt 2520
ttcaaaagag cacccaatca cgttagcaat tggcgatggt gcaaatgatg tcagcatgat 2580
tctggaagcg cacgtgggca taggtgtcat cggcaaggaa ggccgccagg ctgccaggaa 2640
cagcgactat gcaatcccaa agtttaagca tttgaagaag atgctgcttg ttcacgggca 2700
tttttattac attaggatct ctgagctcgt gcagtacttc ttctataaga acgtctgctt 2760
catcttccct cagtttttat accagttctt ctgtgggttt tcacaacaga ctttgtacga 2820
caccgcgtat ctgaccctct acaacatcag cttcacctcc ctccccatcc tcctgtacag 2880
cctcatggag cagcatgttg gcattgacgt gctcaagaga gacccgaccc tgtacaggga 2940
cgtcgccaag aatgccctgc tgcgctggcg cgtgttcatc tactggacgc tcctgggact 3000
gtttgacgca ctggtgttct tctttggtgc ttatttcgtg tttgaaaata caactgtgac 3060
aagcaacggg cagatatttg gaaactggac gtttggaacg ctggtattca ccgtgatggt 3120

gttcacagtt acactaaagc ttgcattgga cacacactac tggacttgga tcaaccattt 3180
tgtcatctgg gggtcgctgc tgttctacgt tgtcttttca cttctctggg gaggagtgat 3240
ctggccgttc ctcaactacc agaggatgta ctacgtgttc atccagatgc tgtccagcgg 3300
gcccgcctgg ctggccatcg tgctgctggt gaccatcagc ctccttcccg acgtcctcaa 3360
gaaagtcctg tgccggcagc tgtggccaac agcaacagag agagtccagc agaatgggtg 3420
cgcacagcct cgggaccgcg actcagaatt caccctctt gcctctctgc agagcccagg 3480
ctaccagagc acctgtccct cggccgcctg gtacagctcc cactctcagc aggtgacact 3540
cgcggcctgg aaggagaagg tgtccacgga gcccccaccc atcctcggcg gttcccatca 3600
ccactgcagt tccatcccaa gtcacagctg ccctaggtcc cgtgtgggaa tgctcgtgtg 3660
atggatggtc ctaagcctgt ggagactgtg cacgtgcctc ttcctggccc ccagcaggca 3720
aggaggggg tcacaggcct tgccctcgaa catggcaccc tggccgcctg gacccagcac 3780
tgt 3783

<210> 50
<211> 2105
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7474127CB1

<400> 50
ccagcgccca gggaagcggc tcaaccacct gaatccggaa aacgccaaca agtagtttct 60
cgtcggagaa gggcggctca cctgggcgcc aagactcagt cccgctgccc agagaacctc 120
gtccactcgg aaaccaaagc agaaccactt ttctctcggt ctcgttaagt catgtctgag 180
tcacagagat gggcaagatc gagaacaacg agagggtgat cctcaatgtc gggggcaccc 240
ggcacgaaac ctaccgcagc accctcaaga ccctgcctgg aacacgcctg gcccttcttg 300
cctcctccga gccccaggc gactgcttga ccacggcggg cgacaagctg cagccgtcgc 360
cgcctccact gtcgccgccg ccgagagcgc ccccgctgtc ccccgggcca ggcggctgct 420
tcgagggcgg cgcgggcaac tgcagttccc gcggcggcag ggccagcgac catcccggtg 480
gcggccgcga gttcttcttc gaccggcacc cgggcgtctt cgcctatgtg ctcaattact 540
accgcaccgg caagctgcac tgccccgcag acgtgtgcgg gccgctcttc gaggaggagc 600
tggccttctg gggcatcgac gagaccgacg tggagccctg ctgctggatg acctaccggc 660
agcaccgcga cgccgaggag gcgctggaca tcttcgagac ccccgacctc attggcggcg 720
accccggcga cgacgaggac ctggcggcca agaggctggg catcgaggac gcggcggggc 780
tcgggggccc cgacggcaaa tctggccgct ggaggaggct gcagccccgc atgtgggccc 840
tcttcgaaga cccctactcg tccagagccg ccaggtttat tgcttttgct tctttattct 900
tcatcctggt ttcaattaca acttttgcc tggaaacaca tgaagctttc aatattgtta 960
aaaacaagac agaaccagtc atcaatggca caagtgttgt tctacaatat gaaattgaaa 1020
cagatcctgc cttgacgtat gtagaaggag tgtgtgtggt gtggtttact tttgaatttt 1080
tagtccgtat tgtttttca cccaataaac ttgaattcat caaaaatctc ttgaatatca 1140
ttgactttgt ggccatccta cctttctact tagaggtggg actcagtggg ctgtcatcca 1200
aagctgctaa agatgtgctt ggcttcctca gggtggtaag gtttgtgagg atcctgagaa 1260
ttttcaagct cacccgccat tttgtaggtc tgagggtgct tggacatact cttcgagcta 1320
gtactaatga atttttgctg ctgataattt tcctggctct aggagttttg atatttgcta 1380
ccatgatcta ctatgccgag agagtgggag ctcaacctaa cgacccttca gctagtgagc 1440
acacacagtt caaaaacatt cccattgggt tctggtgggc tgtagtgacc atgactaccc 1500
tgggttatgg ggatatgtac ccccaaacat ggtcaggcat gctggtggga gccctgtgtg 1560
ctctggctgg agtgctgaca atagccatgc cagtgcctgt cattgtcaat aattttggaa 1620
tgtactactc cttggcaatg gcaaagcaga aacttccaag gaaaagaaag aagcacatcc 1680
ctcctgctcc tcaggcaagc tcacctactt tttgcaagac agaattaaat atggcctgca 1740
atagtacaca gagtgacaca tgtctgggca aagacaatcg acttctggaa cataacagat 1800
cagtgttatc aggtgacgac agtacaggaa gtgagccgcc actatcaccc ccagaaaggc 1860
tccccatcag acgctctagt accagagaca aaaacagaag aggggaaaca tgtttcctac 1920
tgacgacagg tgattacacg tgtgcttctg atggagggat caggaaaggt tatgaaaaat 1980
cccgaagctt aaacaacata gcgggcttgg caggcaatgc tctgaggctc tctccagtaa 2040
catcacccta caactctcct tgtcctctga ggcgctctcg atctcccatc ccatctatct 2100
tgtaa 2105

<210> 51
<211> 2069
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7476949CB1

<400> 51
atgagcaagg acctggcagc aatggggcct ggagcttcag gggacggggt caggactgag 60
acagctccac acatagcact ggactccaga gttggtctgc acgcctacga catcagcgtg 120
gtggtcatct actttgtctt cgtcattgct gtggggatct ggtcgtccat ccgtgcaagt 180
cgagggacca ttggcggcta tttcctggcc gggaggtcca tgagctggtg gccaattgga 240
gcatctctga tgtccagcaa tgtgggcagt ggcttgttca tcggcctggc tgggacaggg 300
gctgccggag gccttgccgt aggtggcttc gagtggaacg caacctggct gctcctggcc 360
cttggctggg tcttcgtccc tgtgtacatc gcagcaggtg tggtcacaat gccgcagtat 420
ctgaagaagc gatttggggg ccagaggatc caggtgtaca tgtctgtcct gtctctcatc 480
ctctacatct tcaccaagat ctcgactgac atcttctctg gagccctctt catccagatg 540
gcattgggct ggaacctgta cctctccaca gggatcctgc tggtggtgac tgccgtctac 600
accattgcag gtggcctcat ggccgtgatc tacacagatg ctctgcagac ggtgatcatg 660
gtaggggggag ccctggtcct catgtttctg ggctttcagg acgtgggctg gtacccaggc 720
ctggagcagc ggtacaggca ggccatccct aatgtcacag tccccaacac cacctgtcac 780
ctcccacggc ccgatgcttt ccacattctt cgggaccctg tgagcgggga catcccttgg 840
ccaggtctca ttttcgggct cacagtgctg gccacctggt gttggtgcac agaccaggtc 900
attgtgcagc ggtctctctc ggccaagagt ctgtctcatg ccaagggagg ctccgtgctg 960
gggggctacc tgaagatcct ccccatgttc ttcatcgtca tgcctggcat gatcagccgg 1020
gccctgttcc cagacgaggt gggctgcgtg gaccctgatg tctgccaaag aatctgtggg 1080
gcccgagtgg gatgttccaa cattgcctac cctaagttgg tcatggccct catgcctgtt 1140
ggtctgcggg ggctgatgat tgccgtgatc atggccgctc tcatgagctc actcacctcc 1200
atcttcaaca gcagcagcac cctgttcacc attgatgtgt ggcagcgctt ccgcaggaag 1260
tcaacagagc aggagctgat ggtggtgggc agagtgtttg tggtgttcct ggttgtcatc 1320
agcatcctct ggatccccat catccaaagc tccaacagtg ggcagctctt cgactacatc 1380
caggctgtca ccagttacct ggccccaccc atcaccgctc tcttcctgct ggccatcttc 1440
tgcaagaggg tcacagagcc cggagctttc tggggcctcg tgtttggcct gggagtgggg 1500
cttctgcgta tgatcctgga gttctcatac ccagcgccag cctgtgggga ggtggaccgg 1560
aggccagcag tgctgaagga cttccactac ctgtactttg caatcctcct ctgcgggctc 1620
actgccatcg tcattgtcat tgtcagcctc tgtacaactc ccatccctga ggaacagctc 1680
acacgcctca catggtggac tcggaactgc cccctctctg agctggagaa ggaggcccac 1740
gagagcacac cggagatatc cgagaggcca gccggggagt gccctgcagg aggtggagcg 1800
gcagagaact cgagcctggg ccaggagcag cctgaagccc caagcaggtc ctggggaaag 1860
ttgctctgga gctggttctg tgggctctct ggaacaccgg agcaggccct gagcccagca 1920
gagaaggctg cgctagaaca gaagctgaca agcattgagg aggagccact ctggagacat 1980
gtctgcaaca tcaatgctgt ccttttgctg gccatcaaca tcttcctctg gggctatttt 2040
gcgtgattca aacctggctt cactgtaga 2069

<210> 52
<211> 4245
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477249CB1

<400> 52
gcggcggcag gctcagctgc gccgggcggg ggcggcgctg gggccgcgcc tgtaggactc 60
ggggccgacg ccgcgggatg gggacgcggc gcggggagtg aggcagtggc ggcggcggcg 120
gtaagcggaa cttcggcccg aggggctcgc ccgctcccgc ctctgtcttg tcggcctcca 180
cctgcagccc cgcggccccc gcgccccgcg ggacccggac ggcgacgacg ggggaatgtg 240
gcgctggatc cggcagcagc tgggttttga cccaccacat cagagtgaca caagaaccat 300
ctacgtagcc aacaggtttc ctcagaatgg cctttacaca cctcagaaat ttatagataa 360
caggatcatt tcatctaagt acactgtgtg gaattttgtt ccaaaaaatt tatttgaaca 420
gttcagaaga gtggcaaact tttatttttct tattatattt ttggttcagc ttatgattga 480
tacacctacc agtccagtta ccagtggact tccattattc tttgtgataa cagtaactgc 540
cataaagcag ggatatgaag attggttacg gcataactca gataatgaag taaatggagc 600
tcctgtttat gttgttcgaa gtggtggcct tgtaaaaact agatcaaaaa acattcgggt 660
gggtgatatt gttcgaatag ccaaagatga aatttttcct gcagacttgg tgcttctgtc 720
ctcagatcga ctggatggtt cctgtcacgt tacaactgct agtttggacg gagaaactaa 780
cctgaagaca catgtggcag ttccagaaac agcattatta caaacagttg ccaatttgga 840
cactctagta gctgtaatag aatgccagca accagaagca gacttataca gattcatggg 900
acgaatgatc ataacccaac aaatggaaga aattgtaaga cctctggggc cggagagtct 960
cctgcttcgt ggagccagat taaaaaacac aaaagaaatt tttggtgttg cggtatacac 1020
tggaatggaa actaagatgg cattaaatta caagagcaaa tcacagaaac gatctgcagt 1080

agaaaagtca atgaatacat ttttgataat ttatctagta attcttatat ctgaagctgt 1140
catcagcact atcttgaagt atacatggca agctgaagaa aaatgggatg aaccttggta 1200
taaccaaaaa acagaacatc aaagaaatag cagtaaggta gagtacgtgt ttacagataa 1260
aactggtaca ctgacagaaa atgagatgca gtttcgggaa tgttcaatta atggcatgaa 1320
ataccaagaa attaatggta gacttgtacc cgaaggacca acaccagact cttcagaagg 1380
aaacttatct tatcttagta gtttatcccá tcttaacaác ttatcccatc ttacaaccag 1440
ttcctctttc agaaccagtc ctgaaaatga aactgaacta attaaagaac atgatctctt 1500
ctttaaagca gtcagtctct gtcacactgt acagattagc aatgttcaaa ctgactgcac 1560
tggtgatggt ccctggcaat ccaacctggc accatcgcag ttggagtact atgcatcttc 1620
accagatgaa aaggctctag tagaagctgc tgcaaggatt ggtattgtgt ttattggcaa 1680
ttctgaagaa actatggagg ttaaaactct tggaaaactg gaacggtaca aactgcttca 1740
tattctggaa tttgattcag atcgtaggag aatgagtgta attgttcagg caccttcagg 1800
tgagaagtta ttatttgcta aaggagctga gtcatcaatt ctccctaaat gtataggtgg 1860
agaaatagaa aaaaccagaa ttcatgtaga tgaatttgct ttgaaagggc taagaactct 1920
gtgtatagca tatagaaaat ttacatcaaa agagtatgag gaaatagata aacgcatatt 1980
tgaagccagg actgccttgc agcagcggga agagaaattg gcagctgttt tccagttcat 2040
agagaaagac ctgatattac ttggagccac agcagtagaa gacagactac aagataaagt 2100
tcgagaaact attgaagcat tgagaatggc tggtatcaaa gtatgggtac ttactgggga 2160
taaacatgaa acagctgtta gtgtgagttt atcatgtggc cattttcata gaaccatgaa 2220
catccttgaa cttataaacc agaaatcaga cagcgagtgt gctgaacaat tgaggcagct 2280
tgccagaaga attacagagg atcatgtgat tcagcatggg ctggtagtgg atgggaccag 2340
cctatctctt gcactcaggg agcatgaaaa actatttatg gaagtttgca gaaattgttc 2400
agctgtatta tgctgtcgta tggctccact gcagaaagca aaagtaataa gactaataaa 2460
aatatcacct gagaaaccta taacattggc tgttggtgat ggtgctaatg acgtaagcat 2520
gatacaagaa gcccatgttg gcataggaat catgggtaaa gaaggaagac aggctgcaag 2580
aaacagtgac tatgcaatag ccagatttaa gttcctctcc aaattgcttt ttgttcatgg 2640
tcatttttat tatattagaa tagctaccct tgtacagtat ttttttata agaatgtgtg 2700
ctttatcaca ccccagtttt tatatcagtt ctactgtttg ttttctcagc aaacattgta 2760
tgacagcgtg tacctgactt tatacaatat ttgttttact tccctaccta ttctgatata 2820
tagtcttttg gaacagcatg tagaccctca tgtgttacaa aataagccca ccctttatcg 2880
agacattagt aaaaaccgcc tcttaagtat taaaacattt ctttattgga ccatcctggg 2940
cttcagtcat gcctttattt tctttttttgg atcctattta ctaataggga aagatacatc 3000
tctgcttgga aatggccaga tgtttggaaa ctggacattt ggcactttgg tcttcacagt 3060
catggttatt acagtcacag taaagatggc tctggaaact catttttgga cttggatcaa 3120
ccatctcgtt acctggggat ctattatatt ttatttttgta ttttccttgt tttatggagg 3180
gattctctgg ccatttttgg gctcccagaa tatgtatttt gtgtttattc agctcctgtc 3240
aagtggttct gcttggtttg ccataatcct catggttgtt acatgtctat ttcttgatat 3300
cataaagaag gtctttgacc gacacctcca ccctacaagt actgaaaagg cacagcttac 3360
tgaaacaaat gcaggtatca agtgcttgga ctccatgtgc tgtttcccgg aaggagaagc 3420
agcgtgtgca tctgttggaa gaatgctgga acgagttata ggaagatgta gtccaaccca 3480
catcagcagg tgtgaaatct ctctaagtag cctttgctgc agatgagtat cctatctgga 3540
acaggatgaa cctgccgctc tagataccta ataaatcagc agctggtttt accaactgaa 3600
gcaggaagtc tgctatttat tagcactctt tggtggtaga tttcactttg tggctttggg 3660
gtaagggctt tttcactcac aaaggaagag aaagcacctt tgaagagact tcatctaatg 3720
aacaaaaaat tttgtttcat aatctttcta aaatgtgctc agtaggagtg tgtttatggt 3780
actcttttat ggtttgtata actttctttt ttaaattata catatactat ttcctttta 3840
ttttttaaa attttttgc tttttgtctt tacaaaataa tctcaacata acagtgaagt 3900
caaaggcttt ccttttctta ctctgtatgt atattttcca gttggttatt tgaggctttg 3960
aggtatttat aaacacaaaa ggctgtattt ctgctcccct acctcttctt atgtctgtaa 4020
tgaagttttg aaatgagtca tgatttttaa gtttcttttg cttggtattt attgcctaat 4080
taaaagtgta tgagttagaa caggcttttt aaattatgga gtaaaagaat cttagcattt 4140
ttgtcccctc ctaaatctgt ttcttgaatg agatttatca ccatgcctgc tgttgtgcac 4200
cataacgaaa aaaaacacct tttggtaaac accatttaaa attca 4245

<210> 53
<211> 2124
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477720CB1

<400> 53
atggctctgc agatgttcgt gacttacagt ccttggaatt gtttgctact gctagtggct 60
cttgagtgtt ctgaagcatc ttctgatttg aatgaatctg caaattccac tgctcagtat 120
gcatctaacg cttggtttgc tgctgccagc tcagagccag aggaagggat atctgttttt 180

gaactggatt atgactatgt gcaaattcct tatgaggtca ctctctggat acttctagca 240
tcccttgcaa aaataggctt ccacctctac cacaggctgc caggcctcat gccagaaagc 300
tgcctcctca tcctggtggg ggcgctggtg ggcggcatca tcttcggcac cgaccacaaa 360
tcgcctccgg tcatggactc cagcatctac ttcctgtatc tcctgccacc catcgttctg 420
gagggcggct acttcatgcc cacqcggccc ttctttgaga acatcggctc catcctgtgg 480
tgggcagtat tgggggccct gatcaacgcc ttgggcattg gcctctccct ctacctcatc 540
tgccaggtga aggcctttgg cctgggcgac gtcaacctgc tgcagaacct gctgttcggc 600
agcctgatct ccgccgtgga cccagtggcc gtgctagccg tgtttgagga agcgcgcgtg 660
aacgagcagc tctacatgat gatctttggg gaggccctgc tcaatgatgg cattactgtg 720
gtcttataca atatgttaat tgcctttaca aagatgcata aatttgaaga catagaaact 780
gtcgacattt tggctggatg tgcccgattc atcgttgtgg ggcttggagg ggtattgttt 840
ggcatcgttt ttggatttat ttctgcattt atcacacgtt tcactcagaa tatctctgca 900
attgagccac tcatcgtctt catgttcagc tatttgtctt acttagctgc tgaaaccctc 960
tatctctccg gcatcctggc aatcacagcc tgcgcagtaa caatgaaaaa gtacgtggaa 1020
gaaaacgtgt cccagacatc atacacgacc atcaagtact tcatgaagat gctgagcagc 1080
gtcagcgaga ccttgatctt catcttcatg ggtgtgtcca ctgtgggcaa gaatcacgag 1140
tggaactggg ccttcatctg cttcaccctg gccttctgcc aaatctggag agccatcagc 1200
gtatttgctc tcttctatat cagtaaccag tttcggactt tccccttctc catcaaggac 1260
cagtgcatca ttttctacag tggtgttcga ggagctggaa gtttttcact tgcatttttg 1320
cttcctctgt ctcttttcc taggaagaaa atgtttgtca ctgctactct agtagttata 1380
tactttactg tatttattca gggaatcaca gttggccctc tggtcaggta cctggatgtt 1440
aaaaaaacca ataaaaaaga atccatcaat gaagagcttc atattcgtct gatggatcac 1500
ttaaaggctg gaatcgaaga tgtgtgtggg cactggagtc actaccaagt gagagacaag 1560
tttaagaagt ttgatcatag atacttacgg aaaatcctca tcagaaagaa cctacccaaa 1620
tcaagcattg tttctttgta caagaagctg gaaatgaagc aagccatcga gatggtggag 1680
actgggatac tgagctctac agctttctcc ataccccatc aggcccagag gatacaagga 1740
atcaaaagac tttcccctga agatgtggag tccataaggg acattctgac atccaacatg 1800
taccaagttc ggcaaaggac cctgtcctac aacaaataca acctcaaacc ccaaacaagt 1860
gagaagcagg ctaaagagat tctgatccgc cgccagaaca ccttaaggga gagcatgagg 1920
aaaggtcaca gcctgccctg gggaaagccg gctggcacca agaatatccg ctacctctcc 1980
taccctacg ggaatcctca gtctgcagga agagacacaa gggctgctgg gttctcaggt 2040
aagctgccca cctggctgct ctgctgcttt tctgtagagt caggtggtaa atatctgggg 2100
gtgtgggcca agaggcaaca ttaa 2124

<210> 54
<211> 2195
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477852CB1

<400> 54
atgggggttt ttctacctaa ggcagaaggg cccgggagcc aactccagaa acttctgccc 60
tcctttctgg tcagagaaca agactgggac cagcacctgg acaagcttca tatgctgcag 120
cagaagagga ttctagagtc tccactgctt cgagcatcca aggaaaatga cctgtctgtt 180
cttaggcaac ttctactgga ctgcacctgt gacgttcgac aaagaggagc cctgggggag 240
acggcgctgc acatagcagc cctctatgac aacttggagg cggccttggt gctgatggag 300
gctgccccag agctggtctt tgagcccacc acatgtgagg cttttgcagg tcagactgca 360
ctgcacatcg ctgttgtgaa ccagaatgtg aacctggtgc gtgccctgct cacccgcagg 420
gccagtgtct ctgccagagc cacaggcact gccttccgcc gtagtccccg caacctcatc 480
tactttggtg agcacccttt gtcctttgct gcctgtgtga acagcgagga gatcgtgcgg 540
ctgctcattg agcatggagc tgacatcagg gcccaggact ccctgggtaa cacagtatta 600
cacatcctca tcctccagcc caacaaaacc tttgcctgcc agatgtacaa cctgctgctg 660
tcctacgatg gacatgggga ccacctgcag cccctggacc ttgtgcccaa tcaccagggt 720
ctcacccctt tcaagctggc tggagtggag ggtaacactg tgatgttcca gcacctgatg 780
cagaagcgga ggcacatcca gtggacgtat ggacccctga cctccattct ctacgacctc 840
acagagatcg actcctgggg agaggagctg tccttcctgg agcttgtggt ctcctctgat 900
aaacgagagg ctcgccaaat tctggaacag accccagtga aggagctggt gagcttcaag 960
tggaacaagt atggccggcc gtacttctgc atcctggctg ccttgtacct gctctacatg 1020
atctgcttta ccacgtgctg cgtctaccgc ccccttaagt ttcgtggtgg caaccgcact 1080
cattctcgag acatcaccat cctccagcaa aaactactac aggaggccta tgagacacgt 1140
gaagatatca tcaggctggt gggggagctg gtgagcatcg ttggggctgt gatcatcctg 1200
ctcctagaga ttccagacat cttcagggtt ggtgcctctc gctattttgg aaagacgatt 1260
cttgggggc cattccatgt catcatgatc acctatgcct ccctggtgct ggtgaccatg 1320
gtgatgcggc tcaccaacac caatggggag gtggtgccca tgtcctttgc cctggtgctg 1380

ggctggtgca gtgtcatgta tttcactcga ggattccaga tgctgggtcc cttcaccatc 1440
atgatccaga agatgatttt tggagaccta atgcgtttct gctggctgat ggctgtggtc 1500
atcttgggat ttgcctccgc gttctatatc attttccaga cagaggaccc aaccagtctg 1560
gggcaattct atgactaccc catggcactg ttcaccacct ttgagctttt tctcactgtt 1620
attgatgcac ctgccaacta cgacgtggac ttgcccttca tgttcagcat tgtcaacttc 1680
gccttcgcca tcattgccac actgctcatg ctcaacttgt tcatcgccat gatgggcgac 1740
acccactgga gggtggccca ggagagggat gagctctgga gggcccaggt cgtggccacc 1800
acagtgatgc tggagcggaa gctgcctcgc tgcctgtggc ctcgctccgg gatctgtggg 1860
tgcgaattcg ggctggggga ccgctggttc ctgcgggttg agaaccacaa tgatcagaat 1920
cctctgcgag tgcttcgcta tgtggaagtg ttcaagaact cagacaagga ggatgaccag 1980
gagcatccat ctgagaaaca gccctctggg gctgagagtg ggactctagc cagagcctct 2040
ttggctcttc caacttcctc cctgtcccgg accgcgtccc agagcagcag tcaccgaggc 2100
tgggagatcc ttcgtcaaaa caccctgggg cacttgaatc ttggactgaa ccttagtgag 2160
ggggatggag aggaggtcta ccatttttga ttaac 2195

<210> 55
<211> 2055
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1471717CB1

<400> 55
cggctctggg ccctcagcct ggctcatgca caactgtctg aagtgctctg gactatggtg 60
atgaacagcg gccttcagac gcgaggctgg ggaggaatcg tcggggtttt tattattttt 120
gccgtatttg ctgtcctgac agtagccatc cttctgatca tggagggcct ctctgctttc 180
ctgcacgccc tgcgactgca ctgggaagct gtttggggaa cttgagctat ttagaagatg 240
gcaaccaagc caacagagcc tgtcacgatc ctcagccttc ggaaattgag cctggggacc 300
gcagagccac aggttaaaga gccaaagacg ttcaccgtgg aagatgcagt ggagactatc 360
ggcttcgggc gtttccacat tgccctcttt ctgatcatgg gcagtactgg ggtggttgag 420
gccatggaga tcatgttgat agctgttgtg tctcctgtca tccgctgtga atggcaactg 480
gagaattggc aggtggcatt agtaaccacg atggtgtttt ttggctacat ggttttcagt 540
atcctctttg gcctcctggc tgacagatat ggccgctgga agattctgct catctcgttc 600
ctgtggggag cctatttctc cttgctgacc tcgtttgctc cttcgtacat ctggtttgtc 660
ttcctgcgga cgatggtggg ctgtggtgtg tccggccact cgcaagggtt aatcataaag 720
actgaatttt tgcccacgaa ataccgaggc tatatgttac ccttgtctca ggtgttctgg 780
cttgcgggct ccctgctcat cattggcttg gcctctgtga tcatccccac catcgggtgg 840
cgctggctca ttcgcgtcgc ctccatcccg ggcatcatcc tcatcgtggc cttcaagttt 900
attcctgaat ctgcccggtt caatgtctcc actgggaaca ctcgggctgc cctggccact 960
ctggagcgcg ttgccaagat gaaccgctcg gtcatgccgg aggggaagct ggtggagccc 1020
gtcctggaaa aaagaggaag atttgcagac ctattggatg ctaaatattt acggaccaca 1080
ttacagatct gggtcatatg gcttggaatc tcttttgcct actatggggt tatcctggcc 1140
agtgctgagc tgctggagcg ggacttggtc tgtggttcaa agtcagactc tgcggtggtg 1200
gtgactgggg gggactcagg ggagagccag agcccctgct actgccacat gtttgcaccc 1260
tctgactatc ggaccatgat catcagcacc atcggtgaaa ttgctttgaa tcctttaaat 1320
atactgggca tcaatttcct gggaagacgg ctgagccttt ctattaccat gggatgcacg 1380
gctttattct gccttctcct caacatttgc acttcaagtg ccggcctgat tggcttcctc 1440
ttcatgctga gggctctggt agctgcaaac ttcaacaccg tctacattta cacagctgag 1500
gtctaccgca ccacgatgcg cgctttgggg atgggaacca gcggctccct gtgtcgcatt 1560
ggtgcaatgg tggcgccatt tatatcccag gttcttatga gtgcatcaat actgggggcc 1620
ctgtgtctct tctcatctgt ctgtgttgta tgcgccattt ctgcattcac tctccccatc 1680
gaaaccaaag gacgggccct ccagcaaatt aaatgaagac ctgcaaagct atgtctacca 1740
gatgagaaaa atgaattcta tcttcagaac tgcggtgcat ttttttaaaa cttggtttta 1800
cttctgtatg ctactcggta attagtaaag tgatttttt ttaaaaggca tatatgggaa 1860
tggggtaggt aactgtatat tgatctcttc cttgaggaac aatatataaa gtacttttat 1920
aaaatataat ttaagctttc aaaggggtgt gagagggaga tggtgggggg gaagatggct 1980
tttcttcgtt gaaatcaagt ctgtaaacct ttatatgaat aaatactaaa ttttaaactt 2040
acaaaaaaaa aaaaa 2055

<210> 56
<211> 4727
<212> DNA
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 3874406CB1

<400> 56
aagagctgct ggagtaggca cccatttaaa gaaaaaatga agaagcagca ataaagaagt 60
tgtaatcgtt acctagacaa acagagaact ggttttgaca gtgtttctag agtgcttttt 120
attattttcc tgacagttgt gttccaccat gattactttc tccttcagcg aataggctaa 180
atgaatatga aacagaaaag cgtgtatcag caaaccaaag cacttctgtg caagaatttt 240
cttaagaaat ggaggatgaa aagagagagc ttattggaat ggggcctctc aatacttcta 300
ggactgtgta ttgctctgtt ttccagttcc atgagaaatg tccagtttcc tggaatggct 360
cctcagaatc tgggaagggt agataaattt aatagctctt ctttaatggt tgtgtataca 420
ccaatatcta atttaaccca gcagataatg aataaaacag cacttgctcc tcttttgaaa 480
ggaacaagtg tcattggggc accaaataaa acacacatgg acgaaatact tctggaaaat 540
ttaccatatg ctatgggaat catctttaat gaaactttct cttataagtt aatatttttc 600
cagggatata acagtccact ttggaaagaa gatttctcag ctcattgctg ggatggatat 660
ggtgagtttt catgtacatt gaccaaatac tggaatagag gatttgtggc tttacaaaca 720
gctattaata ctgccattat agaagtagct ttggtgttcc tgatgagtgt gctgttaaag 780
aaagctgtcc tcaccaattt ggttgtgttt ctccttaccc tcttttgggg atgtctggga 840
ttcactgtat tttatgaaca acttccttca tctctggagt ggattttgaa tatttgtagc 900
ccttttgcct ttactactgg aatgattcag attatcaaac tggattataa cttgaatggt 960
gtaatttttc ctgacccttc aggagactca tatacaatga tagcaacttt ttctatgttg 1020
cttttggatg gtctcatcta cttgctattg gcattatact ttgacaaaat tttaccctat 1080
ggagatgagc gccattattc tcctttattt ttcttgaatt catcatcttg tttccaacac 1140
caaaggacta atgctaaggt tattgagaaa gaaatcgatg ctgagcatcc ctctgatgat 1200
tattttgaac cagtagctcc tgaattccaa ggaaaagaag ccatcagaat cagaaatgtt 1260
aagaaggaat ataaaggaaa atctggaaaa gtggaagcat tgaaaggctt gctctttgac 1320
atatatgaag gtcaaatcac ggcaatcctg ggtcacagtg gagctggcaa atcttcactg 1380
ctaaatattc ttaatggatt gtctgttcca acagaaggat cagttaccat ctataataaa 1440
aatctctctg aaatgcaaga cttggaggaa atcagaaaga taactggcgt ctgtcctcaa 1500
ttcaatgttc aatttgacat actcaccgtg aaggaaaacc tcagcctgtt tgctaaaata 1560
aaagggattc atctaaagga agtggaacaa gaggtacaac gaatattatt ggaattggac 1620
atgcaaaaca ttcaagataa ccttgctaaa catttaagtg aaggacagaa aagaaagctg 1680
acttttggga ttaccatttt aggagatcct caaattttgc tcttagatga accaactact 1740
ggattggatc ccttttccag agatcaagtg tggagcctcc tgagagagcg tagagcagat 1800
catgtgatcc ttttcagtac ccagtccatg gatgaggctg acatcctggc tgatagaaaa 1860
gtgatcatgt ccaatgggag actgaagtgt gcaggttctt ctatcttttt gaaaagaagg 1920
tggggtcttg gatatcacct aagtttacat aggaatgaaa tatgtaaccc agaacaaata 1980
acatccttca ttactcatca catccccgat gctaaattaa aaacagaaaa caaagaaaag 2040
cttgtatata ctttgccact ggaaaggaca aatacatttc cagatctttt cagtgatctg 2100
gataagtgtt ctgaccaggg agtgacaggt tatgacattt ccatgtcaac tctaaatgaa 2160
gtctttatga aactggaagg acagtcaact atcgaacaag atttcgaaca agtggagatg 2220
ataagagact cagaaagcct caatgaaatg gagctggctc actcttcctt ctctgaaatg 2280
cagacagctg tgagtgacat gggcctctgg agaatgcaag tctttgccat ggcacggctc 2340
cgtttcttaa agttaaaacg tcaaactaaa gtgttattga ccctattatt ggtatttgga 2400
atcgcaatat tccctttgat tgttgaaaat ataatatatg ctatgttaaa tgaaaagatc 2460
gattgggaat ttaaaaacga attgtatttt ctctctcctg gacaacttcc ccaggaaccc 2520
cgtaccagcc tgttgatcat caataacaca gaatcaaata ttgaagattt tataaaatca 2580
ctgaagcatc aaaatatact tttggaagta gatgactttg aaaacagaaa tggtactgat 2640
ggcctctcat acaatggagc tatcatagtt tctggtaaac aaaaggatta tagattttca 2700
gttgtgtgta ataccaagag attgcactgt tttccaattc ttatgaatat tatcagcaat 2760
gggctacttc aaatgtttaa tcacacacaa catattcgaa ttgagtcaag cccatttcct 2820
cttagccaca taggactctg gactgggttg ccggatggtt ccttttctt attttggtt 2880
ctatgtagca tttctcctta tatcaccatg ggcagcatca gtgattacaa gaaaaatgct 2940
aagtcccagc tatggatttc aggcctctac acttctgctt actggtgtgg gcaggcacta 3000
gtggacgtca gcttcttcat tttaattctc cttttaatgt atttaatttt ctacatagaa 3060
aacatgcagt accttcttat tacaagccaa attgtgtttg ctttggttat agttactcct 3120
ggttatgcag cttctcttgt cttcttcata tatatgatat catttatttt tcgcaaaagg 3180
agaaaaaaca gtggcctttg gtcattttac ttctttttg cctccaccat catgttttcc 3240
atcactttaa tcaatcattt tgacctaagt atattgatta ccaccatggt attggttcct 3300
tcatatacct tgcttggatt taaaactttt ttggaagtga gagaccagga gcactacaga 3360
gaatttccag aggcaaattt tgaattgagt gccactgatt ttctagtctg cttcataccc 3420
tactttcaga ctttgctatt cgtttttgtt ctaagatgca tggaactaaa atgtggaaag 3480
aaaagaatgc gaaaagatcc tgttttcaga atttcccccc aaagtagaga tgctaagcca 3540
aatccagaag aacccataga tgaagatgaa gatattcaaa cagaaagaat aagaacagtc 3600
actgctctga ccacttcaat cttagatgag aaacctgtta taattgccag ctgtctcacac 3660
aaagaatatg caggccagaa gaaaagttgc ttttcaaaga ggaagaagaa aatagcagca 3720
agaaatatct ctttctgtgt tcaagaaggt gagattttgg gattgctagg acccagtggt 3780

gctggaaaaa gttcatctat tagaatgata tctgggatca caaagccaac tgctggagag 3840
gtggaactga aaggctgcag ttcagttttg ggccacctgg ggtactgccc tcaagagaac 3900
gtgctgtggc ccatgctgac gttgagggaa cacctggagg tgtatgctgc cgtcaagggg 3960
ctcagggaag cggacgcgag gctcgccatc gcaagattag tgagtgcttt caaactgcat 4020
gagcagctga atgttcctgt gcagaaatta acagcaggaa tcacgagaaa gttgtgtttt 4080
gtgctgagcc tcctgggaaa ctcacctgtc ttgctcctgg atgaaccatc tacgggcata 4140
gaccccacag ggcagcagca aatgtggcag gcaatccagg cagtcgttaa aaacacagag 4200
agaggtgtcc tcctgaccac ccataacctg gctgaggcgg aagccttgtg tgaccgtgtg 4260
gccatcatgg tgtctggaag gcttagatgc attggctcca tccaacacct gaaaaacaaa 4320
cttggcaagg attacattct agagctaaaa gtgaaggaaa cgtctcaagt gactttggtc 4380
cacactgaga ttctgaagct tttcccacag gctgcagggc agcaaaggta ttcctctttg 4440
ttaacctata agctgcccgt ggcagacgtt taccctctat cacagacctt tcacaaatta 4500
gaagcagtga agcataactt taacctggaa gaatacagcc tttctcagtg cacactggag 4560
aaggtattct tagagctttc taaagaacag gaagtaggaa attttgatga agaaattgat 4620
acaacaatga gatggaaact cctcctcat tcagatgaac cttaaaacct caaacctagt 4680
aatttttcttg cttgatctcc tataaactta tgttttatgt aataatt 4727

<210> 57
<211> 3852
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4599654CB1

<400> 57
cgccggcgat tccgagccta cgacgcctcc gctagagccc gcggggctgc gccgactcct 60
gctctggagg ggttgcgggt acctgatggc cacagagggc tctaggaggc cgagcgtgta 120
agcggggtgg gcgccatgga ggcagagcag cggccggcgg cgggggccag cgaaggggcg 180
acccctggac tggaggcggt gcctcccgtt gctcccccgc ctgcgaccgc ggcctcaggt 240
ccgatcccca aatctgggcc tgagcctaag aggaggcacc ttgggacgct gctccagcct 300
acggtcaaca agtttctccct tcgggtgttc ggcagccaca aagcagtgga aatcgagcag 360
gagcgggtga agtcagcggg ggcctggatc atccacccct acagcgactt ccggttttac 420
tgggacctga tcatgctgct gctgatggtg gggaacctca tcgtcctgcc tgtgggcatc 480
accttcttca aggaggagaa ctccccgcct tggatcgtct tcaacgtatt gtctgatact 540
ttcttcctac tggatctggt gctcaacttc cgaacgggca tcgtggtgga ggagggtgct 600
gagatcctgc tggcaccgcg ggccatccgc acgcgctacc tgcgcacctg gttcctggtt 660
gacctcatct cttctatccc tgtggattac atcttcctag tggtggagct ggagccacgg 720
ttggacgctg aggtctacaa aacggcacgg gccctacgca tcgttcgctt caccaagatc 780
ctaagcctgc tgaggctgct ccgcctctcc cgcctcatcc gctacataca ccagtgggag 840
gagatcttcc acatgaccta tgacctggcc agtgctgtgg ttcgcatctt caacctcatt 900
gggatgatgc tgctgctatg tcactgggat ggctgtctgc agttcctggt gcccatgctg 960
caggacttcc ctcccgactg ctgggtctcc atcaaccaca tggtgaacca ctcgtggggc 1020
cgccagtatt cccatgccct gttcaaggcc atgagccaca tgctgtgcat tggctatggg 1080
cagcaggcac ctgtaggcat gcccgacgtc tggctcacca tgctcagcat gatcgtaggt 1140
gccacatgct acgccatgtt catcggccat gccacggcac tcatccagtc cctggactct 1200
tcccggcgtc agtaccagga gaagtacaag caggtggagc agtacatgtc cttccacaag 1260
ctgccagcag acacgcggca gcgcatccac gagtactatg agcaccgcta ccagggcaag 1320
atgttcgatg aggaaagcat cctgggcgag ctgagcgagc cgcttcgcga ggagatcatt 1380
aacttcacct gtcggggcct ggtggcccac atgccgctgt ttgcccatgc cgaccccagc 1440
ttcgtcactg cagttctcac caagctgcgc tttgaggtct tccagccggg ggatctcgtg 1500
gtgcgtgagg gctccgtggg gaggaagatg tacttcatcc agcatgggct gctcagtgtg 1560
ctggcccgcg gcgcccggga cacacgcctc accgatggat cctactttgg ggagatctgc 1620
ctgctaacta ggggccggcg cacagccagt gttcgggctg acacctactg ccgcctttac 1680
tcactcagcg tggaccattt caatgctgtg cttgaggagt tccccatgat gcgccgggcc 1740
tttgagactg tggccatgga tcggctgctc cgcatcggca agaagaattc catactgcag 1800
cggaagcgct ccgagccaag tccaggcagc agtggtggca tcatggagca gcatttggtg 1860
caacatgaca gagacatggc tcggggtgtt cggggtcggg ccccgagcac aggagctcag 1920
cttagtggaa agccagtact gtgggagcca ctggtacatg cgccccttca ggcagctgct 1980
gtgacctcca atgtggccat tgccctgact catcagcggg gccctctgcc cctctccccct 2040
gactctccag ccaccctcct tgctcgctct gcttggcgct cagcaggctc tccagcttcc 2100
ccgctggtgc ccgtccgagc tggcccatgg gcatccacct cccgcctgcc cgccccacct 2160
gcccgaaccc tgcacgccag cctatcccgg gcagggcgct cccaggtctc cctgctgggt 2220
ccccctccag gaggaggtgg acggcggcta ggacctcggg gccgcccact ctcagcctcc 2280
caaccctctc tgcctcagcg ggcaacaggc gatggctctc ctgggcgtaa gggatcagga 2340
agtgagcggc tgcctccctc agggctcctg gccaaacctc caaggacagc ccagcccccc 2400

aggccaccag tgcctgagcc agccacaccc cggggtctcc agctttctgc caacatgtaa 2460
aacctttgag tacatccagc cttagttctt ggggtgcagt agtatgtacc caagggcaga 2520
tgcctcttgg ggaaggccat gggacctga aacattgccc catggaaatg tcgaccctgt 2580
gcggacattc cgcatactgc catgaagacg gtctctgtgt cctcagctca agaatcctgt 2640
agcttgtccc atcataatcc attcacccgt tcatcatgtg tactgagcag ctaccatgtt 2700
caaggtaata tgccaggcgc tgtatgtctc cactgccaag tagaagtgac tcaaaaccct 2760
ctgacaagga tattcccttg gctatggtcc tgccaggtgc aggcccaggc ccatgacccc 2820
acctttacta agcacaagta cttgccactg ccatcactgc caagtaacta gatgtctctg 2880
tttcctgcc aatgatcctg caggttctgc ccggtctggt tatcttcctg tttcctgtag 2940
catagccagg cactgccagt cacctgtgcc cccattgctg tcagcagatg tcttgggtcc 3000
tgagtgtggg tatccacttt tacccgctca ctgccacctg tggacactct gtgtctaccc 3060
tctgagtggg aacatacttc taagttccct gcagtctctg tcctgtggta gaccatcttt 3120
ttgtaaactg cgagcttcct cttccctgta ccctctgccc cagtcgtgac cccctaaaag 3180
ttaaggggta gttggcacct ccttattaat atgccagcct agatcccccc cggtggaggg 3240
gcaaatggct gaatccttgt gtgatatttt tttcttcgct tgtttattta ttcatttatt 3300
taattgtatt tattcattta ctaactttat gtgttaccaa ttaattttgt ttacccattc 3360
ctttatccat ccctcccctc cttttcaggt aaggagacag gaggagtagg aggaggcagg 3420
gcctctccat gccagcctct gtggtccttg cccaaaccca tcagcgcaat acttgaacct 3480
tctcccaggt aggggcagga ggagccacat gagagaggga gaaggaccgc gtttaccttt 3540
agagttttgt tttgtttttt ccttctgagt ttgctgttgg tgcaggaata agggaaaggc 3600
ccaaggtatc caagcctggg gaagggcagg ccagccagca cctctgcctt ctcagggaca 3660
agagtagtcc tttaccaccc tcactctgcc tgtcccctct cctactctac agcattaaag 3720
actgtgggac caggacccta agtctccttt ccttctgggt ggggagttct aggggttctt 3780
ggtgtgtggg agaagtttta taattgcttc caaacagctg ggtttaaata taaaatagac 3840
acactcaaaa aa 3852

<210> 58
<211> 1917
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5047435CB1

<400> 58
atggcagaag gtgaaagggg agcagacgtg ccacatggcc tcggggcctg gctggccgac 60
gtggcgttgg cggcgctgcg cgcgggaggg cagggcagga gggacagagg cgggggcggg 120
ccggaaagtt tgtccggcgg cagcggcgtt ggggactccg gcgggggatg cgcgcccggc 180
ccctcagcgc ccccagcacg ccgccgagtc ccgctcgcca tgggccactc cccacctgtc 240
ctgcctttgt gtgcctctgt gtctttgctg ggtggcctga cctttggtta tgaactggca 300
gtcatatcag gtgccctgct gccactgcag cttgactttg ggctaagctg cttggagcag 360
gagttcctgg tgggcagcct gctcctgggg gctctcctcg cctccctggt tggtggcttc 420
ctcattgact gctatggcag gaagcaagcc atcctcggga gcaacttggt gctgctggca 480
ggcagcctga ccctgggcct ggctggttcc ctggcctggc tggtcctggg ccgcgctgtg 540
gttggcttcg ccatttccct ctcctccatg gcttgctgta tctacgtgtc agagctggtg 600
gggccacggc agcggggagt gctggtgtcc ctctatgagg caggcatcac cgtgggcatc 660
ctgctctcct atgccctcaa ctatgcactg gctggtaccc cctggggatg gaggcacatg 720
ttcggctggg ccactgcacc tgctgtcctg caatccctca gcctcctctt cctccctgct 780
ggtacagatg agactgcaac acacaaggac ctcatcccac tccagggagg tgaggccccc 840
aagctgggcc cggggaggcc acggtactcc tttctggacc tcttcagggc acgcgataac 900
atgcgaggcc ggaccacagt gggcctgggg ctggtgctct tccagcaact aacagggcag 960
cccaacgtgc tgtgctatgc ctccaccatc ttcagctccg ttggtttcca tgggggatcc 1020
tcagccgtgc tggcctctgt ggggcttggc gcagtgaagg tggcagctac cctgaccgcc 1080
atggggctgg tggaccgtgc aggccgcagg gctctgttgc tagctggctg tgccctcatg 1140
gccctgtccg tcagtggcat aggcctcgtc agctttgccg tgcccatgga ctcaggccca 1200
agctgtctgg ctgtgcccaa tgccaccggg cagacaggcc tccctggaga ctctggcctg 1260
ctgcaggact cctctctacc tcccattcca aggaccaatg aggaccaaag ggagccaatc 1320
ttgtccactg ctaagaaaac caagccccat cccagatctg gagacccctc agcccctcct 1380
cggctggccc tgagctctgc cctccctggg ccccctctgc ccgctcgggg gcatgcactg 1440
ctgcgctgga ccgcactgct gtgcctgatg gtctttgtca gtgccttctc ctttgggttt 1500
gggccagtga cctggcttgt cctcagcgag atctaccctg tggagatacg aggaagagcc 1560
ttcgccttct gcaacagctt caactgggcg gccaacctct tcatcagcct ctccttcctc 1620
gatctcattg gcaccatcgg cttgtcctgg accttcctgc tctacggact gaccgctgtc 1680
ctcggcctgg gcttcatcta tttatttgtt cctgaaacaa aaggccagtc gttggcagag 1740
atagaccagc agttccagaa gagacggttc accctgagct ttggccacag gcagaactcc 1800
actggcatcc cgtacagccg catcgagatc tctgcggcct cctgaggaat ccgtctgcct 1860

ggaaattctg gaactgtggc tttggcagac catctccagc atcctgcttc ctaggcc 1917

<210> 59
<211> 6791
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7475603CB1

<400> 59
cgcgctccct gcctgctgct gggcggaggg aaggcggcaa gagctgcgga gcccctggaa 60
gagcttccag gaaccctgcg ctgtgggata aaggaatgag gttcagaaag gggcaggagt 120
tgcccgcagc cgcaccgcac gtcttcagcc cgaccgttgt cctgacctct ctgtcccgtc 180
ccctgcccag tctcaccatg gccttctgga cacagctgat gctgctgctc tggaagaatt 240
tcatgtatcg ccggagacag ccggtccagc tcctggtcga attgctgtgg cctctcttcc 300
tcttcttcat cctggtggct gttcgccact cccacccgcc cctggagcac catgaatgcc 360
acttcccaaa caagccactg ccatcggcgg gcaccgtgcc ctggctccag ggtctcatct 420
gtaatgtgaa caacacctgc tttccgcagc tgacaccggg cgaggagccc gggcgcctga 480
gcaacttcaa cgactccctg gtctcccggc tgctagccga tgcccgcact gtgctgggag 540
gggccagtgc ccacaggacg ctggctggcc tagggaagct gatcgccacg ctgagggctg 600
cacgcagcac ggcccagcct caaccaacca agcagtctcc actggaacca cccatgctgg 660
atgtcgcgga gctgctgacg tcactgctgc gcacggaatc cctggggttg gcactgggcc 720
aagcccagga gcccttgcac agcttgttgg aggccgctga ggacctggcc caggagctcc 780
tggcgctgcg cagcctggtg gagcttcggg cactgctgca gagaccccga gggaccagcg 840
gccccctgga gttgctgtca gaggccctct gcagtgtcag gggacctagc agcacagtgg 900
gcccctccct caactggtac gaggctagtg acctgatgga gctggtgggg caggagccag 960
aatccgccct gccagacagc agcctgagcc ccgcctgctc ggagctgatt ggagccctgg 1020
acagccaccc gctgtcccgc ctgctctgga gacgcctgaa gcctctgatc ctcgggaagc 1080
tactctttgc accagataca ccttttaccc ggaagctcat ggcccaggtg aaccggacct 1140
tcgaggagct caccctgctg agggatgtcc gggaggtgtg ggagatgctg ggaccccgga 1200
tcttcacctt catgaacgac agttccaatg tggccatgct gcagcggctc ctgcagatgc 1260
aggatgaagg aagaaggcag cccagacctg gaggccggga ccacatggag gccctgcgat 1320
cctttctgga ccctgggagc ggtggctaca gctggcagga cgcacacgct gatgtggggc 1380
acctggtggg cacgctgggc cgagtgacgg agtgcctgtc cttggacaag ctggaggcgg 1440
caccctcaga ggcagccctg gtgtcgcggg ccctgcaact gctcgcggaa catcgattct 1500
gggccggcgt cgtcttcttg ggacctgagg actcttcaga ccccacagag cacccaaccc 1560
cagacctggg ccccggccac gtgcgcatca aaatccgcat ggacattgac gtggtcacga 1620
ggaccaataa gatcagggac aggttttggg accctggccc agccgcggac cccctgaccg 1680
acctgcgcta cgtgtggggc ggcttcgtgt acctgcaaga cctggtggag cgtgcagccg 1740
tccgcgtgct cagcggcgcc aacccccggg ccggcctcta cctgcagcag atgccctatc 1800
cgtgctatgt ggacgacgtg ttcctgcgtg tgctgagccg gtcgctgccg ctcttcctga 1860
cgctggcctg gatctactcc gtgacactga cagtgaaggc cgtggtgcgg gagaaggaga 1920
cgcggctgcg ggacaccatg cgcgccatgg ggctcagccg cgcggtgctc tggctaggct 1980
ggttcctcag ctgcctcggg cccttcctgc tcagcgccgc gctgctggtt ctggtgctca 2040
agctgggggа catcctcccc tacagccacc cgggcgtggt cttcctgttc ttggcagcct 2100
tcgcggtggc cacggtgacc cagagcttcc tgctcagcgc cttcttctcc cgcgccaacc 2160
tggctgcggc ctgcggcggc ctggcctact tctccctcta cctgccctac gtgctgtgtg 2220
tggcttggcg ggaccggctg cccgcgggtg gccgcgtggc cgcgagcctg ctgtcgcccg 2280
tggccttcgg cttcggctgc gagagcctgg ctctgctgga ggagcagggc gagggcgcgc 2340
agtggcacaa cgtgggcacc cggcctacgg cagacgtctt cagcctggcc caggtctctg 2400
gccttctgct gctggacgcg gcgctctacg gcctcgccac ctggtacctg gaagctgtgt 2460
gcccaggcca gtacgggatc cctgaaccat ggaattttcc ttttcggagg agctactggt 2520
gcggacctcg gccccccaag agtccagccc cttgccccac cccgctggac ccaaaggtgc 2580
tggtagaaga ggcaccgccc ggcctgagtc ctggcgtctc cgttcgcagc ctggagaagc 2640
gctttcctgg aagcccgcag ccagccctgc gggggctcag cctggacttc taccagggcc 2700
acatcaccgc cttcctgggc cacaacgggg ccggcaagac caccaccctg tccatcttga 2760
gtggcctctt cccacccagt ggtggctctg ccttcatcct gggccacgac gtccgctcca 2820
gcatggccgc catccggccc cacctgggcg tctgtcctca gtacaacgtg ctgtttgaca 2880
tgctgaccgt ggacgagcac gtctggttct atgggcggct gaagggtctg agtgccgctg 2940
tagtgggccc cgagcaggac cgtctgctgc aggatgtggg gctggtctcc aagcagagtg 3000
tgcagactcg ccacctctct ggtgggatgc aacggaagct gtccgtggcc attgcctttg 3060
tgggcggctc ccaagttgtt atcctggacg agcctacggc tggcgtggat cctgcttccc 3120
gccgcggtat ttgggagctg ctgctcaaat accgagaagg tcgcacgctg atcctctcca 3180
cccaccacct ggatgaggca gagctgctgg gagaccgtgt ggccgtggtg gcaggtggcc 3240
gcttgtgctg ctgtggatcc ccactcttcc tgcgccgtca cctgggctcc ggctactacc 3300

tgacgctggt gaaggcccgc ctgcccctga ccaccaatga gaaggctgac actgacatgg 3360
agggcagtgt ggacaccagg caggaaaaga agaatggcag ccagggcagc agagtcggca 3420
ctcctcagct gctggccctg gtacagcact gggtgcccgg ggcacggctg gtggaggagc 3480
tgccacacga gctggtgctg gtgctgccct acacgggtgc ccatgacggc agcttcgcca 3540
cactcttccg agagctagac acgcggctgg cggagctgag gctcactggc tacgggatct 3600
ccgacaccag cctcgaggag atcttcctga aggtggtgga ggagtgtgct gcggacacag 3660
atatggagga tggcagctgc gggcagcacc tatgcacagg cattgctggc ctagacgtaa 3720
ccctacggct caagatgccg ccacaggaga cagcgctgga gaacgggaaa ccagctgggt 3780
cagccccaga gactgaccag ggctctgggc cagacgccgt gggccgggta cagggctggg 3840
cactgacccg ccagcagctc caggccctgc ttctcaagcg ctttctgctt gcccgccgca 3900
gccgccgcgg cctgttcgcc cagatcgtgc tgcctgccct ctttgtgggc ctggccctcg 3960
tgttcagcct catcgtgcct cctttcgggc actacccggc tctgcggctc agtcccacca 4020
tgtacggtgc tcaggtgtcc ttcttcagtg aggacgcccc aggggaccct ggacgtgccc 4080
ggctgctcga ggcgctgctg caggaggcag gactggagga gccccaagtg cagcatagct 4140
cccacaggtt ctcggcacca gaagttcctg ctgaagtggc caaggtcttg gccagtggca 4200
actggacccc agagtctcca tccccagcct gccagtgtag ccggcccggt gcccggcgcc 4260
tgctgcccga ctgcccggct gcagctggtg gtccccctcc gccccaggca gtgaccggct 4320
ctggggaagt ggttcagaac cagacaggcc ggaacctgtc tgacttcctg gtcaagacct 4380
acccgcgcct ggtgcgccag ggcctgaaga ctaagaagtg ggtgaatgag gtcagatacg 4440
gaggcttctc gctggggggc cgagaaccag gcctgccctc gggccaagag ttgggccgct 4500
cagtggagga gttgtgggcg ctgctgagtc ccctgcctgg cggggccctc gaccgtgtcc 4560
tgaaaaacct cacagcctgg gctcacagcc tggatgctca ggacagtctc aagatctggt 4620
tcaacaacaa aggctggcac tccatggtgg cctttgtcaa ccgagccagc aacgcaatcc 4680
tccgtgctca cctgccccca ggcccggccc gccacgccca cagcatcacc acactcaacc 4740
acccttgaa cctcaccaag gagcagctgt ctgaggctgc actgatggcc tcctcggtgg 4800
acgtcctcgt ctccatctgt gtggtctttg ccatgtcctt tgtcccggcc agcttcactc 4860
ttgtcctcat tgaggagcga gtcacccgag ccaagcacct gcagctcatg gggggcctgt 4920
ccccacccct ctactggctt ggcaactttc tctgggacat gtgtaactac ttggtgccag 4980
catgcatcgt ggtgctcatc tttctggcct tccagcagag ggcatatgtg gcccctgcca 5040
acctgcctgc tctcctgctg ttgctactac tgtatggctg gtcgatcaca ccgctcatgt 5100
acccagcctc cttcttcttc tccgtgccca gcacagccta tgtggtgctc acctgcataa 5160
acctctttat tggcatcaat ggaagcatgg ccacctttgt gcttgagctc ttctctgatc 5220
agaagctgca ggaggtgagc cggatcttga aacaggtctt ccttatcttc ccccacttct 5280
gcttgggccg ggggctcatt gacatggtgc ggaaccaggc catggctgat gcctttgagc 5340
gcttgggaga caggcagttc cagtcacccc tgcgctggga ggtggtcggc aagaacctct 5400
tggccatggt gatacagggg cccctcttcc ttctcttcac actactgctg cagcaccgaa 5460
gccaactcct gccacagccc agggtgaggt ctctgccact cctgggagag gaggacgagg 5520
atgtagcccg tgaacgggag cgggtggtcc aaggagccac ccaggggat gtgttggtgc 5580
tgaggaactt gaccaaggta taccgtgggc agaggatgcc agctgttgac cgcttgtgcc 5640
tggggattcc ccctggtgag tgttttgggc tgctgggtgt gaatggagca gggaagacgt 5700
ccacgtttcg catggtgacg ggggacacat tggccagcag gggcgaggct gtgctggcag 5760
gccacagcgt ggcccgggaa cccagtgctg cgcacctcag catgggatac tgccctcaat 5820
ccgatgccat ctttgagctg ctgacgggcc gcgagcacct ggagctgctt gcgcgcctgc 5880
gcggtgtccc ggaggcccag gttgcccaga ccgctggctc gggcctggcg cgtctgggac 5940
tctcatggta cgcagaccgg cctgcaggca cctacagcgg agggaacaaa cgcaagctgg 6000
cgacggccct ggcgctggtt ggggacccag ccgtggtgtt tctggacgag ccgaccacag 6060
gcatggaccc cagcgcgcgg cgcttccttt ggaacagcct tttggccgtg gtgcgggagg 6120
gccgttcagt gatgctcacc tcccatagca tggaggagtg tgaagcgctc tgctcgcgcc 6180
tagccatcat ggtgaatggg cggttccgct gcctgggcag cccgcaacat ctcaagggca 6240
gattcgcggc gggtcacaca ctgaccctgc gggtgcccgc cgcaaggtcc cagccggcag 6300
cggccttcgt ggcggccgag ttccctgggg cggagctgcg cgaggcacat ggaggccgcc 6360
tgcgcttcca gctgccgccg ggagggcgct gcgccctggc gcgcgtcttt ggagagctgg 6420
cggtgcacgg cgcagagcac ggcgtggagg acttttccgt gagccagacg atgctggagg 6480
aggtattctt gtacttctcc aaggaccagg ggaaggacga ggacaccgaa gagcagaagg 6540
aggcaggagt gggagtggac cccgcgccag gcctgcagca ccccaaacgc gtcagccagt 6600
tcctcgatga ccctagcact gccgagactg tgctctgagc ctccctcccc tgcggggccg 6660
cggggaggcc ctgggaatgg caagggcaag gtagagtgcc taggagccct ggactcaggc 6720
tggcagaggg gctggtgccc tggagaaaat aaagagaagg ctggagagaa gccgtggtgg 6780
tgaaaaaaaa a 6791

<210> 60
<211> 5214
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 7477845CB1

<400> 60

atgctcaaaa ggaagcagag ttccagggtg gaagcccagc cagtcactga ctttggtcct 60
gatgagtctc tgtcggataa tgctgacatc ctctggatta acaaaccatg ggttcactct 120
ttgctgcgca tctgtgccat catcagcgtc atttctgttt gtatgaatac gccaatgacc 180
ttcgagcact atcctccact tcagtatgtg accttcactt tggatacatt attgatgttt 240
ctctacacgg cagagatgat agcaaaaatg cacatccggg gcattgtcaa gggggatagt 300
tcctatgtga aagatcgctg gtgtgttttt gatggattta tggtcttttg cctttgggtt 360
tctttggtgc tacaggtgtt tgaaattgct gatatagttg atcagatgtc accttggggc 420
atgttgcgga ttccacggcc actgattatg atccgagcat tccggattta tttccgattt 480
gaactgccaa ggaccagaat tacaaatatt ttaaagcgat cgggagaaca aatatggagt 540
gtttccattt ttctactttt ctttctactt ctttatggaa ttttaggagt tcagatgttt 600
ggaacattta cttatcactg tgttgtaaat gacacaaagc cagggaatgt aacctggaat 660
agtttagcta ttccagacac acactgctca ccagagctag aagaaggcta ccagtgccca 720
cctggattta aatgcatgga ccttgaagat ctgggactta gcaggcaaga gctgggctac 780
agtggcttta atgagatagg aactagtata ttcaccgtct atgaggccgc ctcacaggaa 840
ggctgggtgt tcctcatgta cagagcaatt gacagctttc cccgttggcg ttcctacttc 900
tatttcatca ctctcatttt cttcctcgcc tggcttgtga agaacgtgtt tattgctgtt 960
atcattgaaa catttgcaga aatcagagta cagtttcaac aaatgtgggg atcgagaagc 1020
agcactaccc caacagccac cacccagatg tttcatgaag atgctgctgg aggttggcag 1080
ctggtagctg tggatgtcaa caagccccag ggacgcgccc cagcctgcct ccagaaaatg 1140
atgcggtcat ccgttttcca catgttcatc ctgagcatgg tgaccgtgga cgtgatcgtg 1200
gcggctagca actactacaa aggagaaaac ttcaggaggc agtacgacga gttctacctg 1260
gcggaggtgg cttttacagt acttttgat ttggaagcac ttctgaagat atggtgtttg 1320
ggatttactg gatatattag ctcatctctc cacaaattcg aactactact cgtaattgga 1380
actactcttc atgtataccc agatctttat cattcacaat tcacgtactt tcaggtactc 1440
cgagtagttc ggctgattaa gatttcacct gcattagaag actttgtgta caagatattt 1500
ggtcctggaa aaaagcttgg gagtttggtt gtatttactg ccagcctctt gattgttatg 1560
tcagcaatta gtttgcagat gttctgcttt gtcgaagaac tggacagatt tactacgttt 1620
ccgagggcat ttatgtccat gttccagatc ctcacccagg aaggatgggt ggacgtaatg 1680
gaccaaactc taaatgctgt gggacatatg tgggcacccg tggttgccat ctatttcatt 1740
ctctatcatc tttttgccac tctgatcctc ctgagtttgt ttgttgctgt tattttggac 1800
aacttagaac ttgatgaaga cctaaagaag cttaaacaat taaagcaaag tgaagcaaat 1860
gcggacacca aagaaaagct ccctttacgc ctgcgaatct ttgaaaaatt tccaaacaga 1920
cctcaaatgg tgaaaatctc aaagcttcct tcagatttta cagttcctaa aatcagggag 1980
agttttatga agcagtttat tgaccgccag caacaggaca catgttgcct tctgagaagc 2040
ctcccgacca cctcttcctc ctcctgcgac cactccaaac gctcagcaat tgaggacaac 2100
aaatacatcg accaaaaact tcgcaagtct gttttcagca tcagggcaag gaaccttctg 2160
gaaaaggaga ccgcagtcac taaaatctta agggcttgca cccgacagcg catgctgagc 2220
ggatcatttg aggggcagcc cgcaaaggag aggtcaatcc tcagcgtgca gcatcatatc 2280
cgccaagagc gcaggtcact aagacatgga tcaaacagcc agaggatcag caggggaaaa 2340
tctcttgaaa ctttgactca agatcattgc aatacagtga tatatagaaa tgctcaaaga 2400
gaagtcagtg aaataaagat gattcaggaa aaaaaggagc tagcagagat gcttcaagga 2460
aagtgcaaaa aggaactcag agagagccac ccatacttcg ataagccact gttcattgtc 2520
gggcgagaac acaggttcag aaacttttgc cgggtggtgg tccgagcacg cttcaacgcg 2580
tctaaaacag accctgtcac aggagctgtg aaaaatacaa agtaccatct tctttatgat 2640
ttgctgggat tggtcactta cctggactgg gtcatgatca tcgtaacctc tgactcttgc 2700
atttccatga tgtttgagtc cccgtttcga agagtcatgc atgcacctac tttgcagatt 2760
gctgagtatg tgtttgtgat attcatgagc attgagctta atctgaagat tatggcagat 2820
ggcttatttt tcactccaac tgctgtcatc agggacttcg gtggagtaat ggacatattt 2880
atatatcttg tgagcttgat atttctttgt tggatgcctc aaaatgtacc tgctgaatcg 2940
ggagctcagc ttctaatggt ccttcggtgc ctgagacctc tgcgcatatt caaactggtg 3000
ccccagatga ggaaagttgt tcgagaactt ttcagcggct tcaaggaaat ttttttggtc 3060
tccattcttt tgctgacatt aatgctcgtt tttgcaagct ttggagttca gctttttgct 3120
ggaaaactgg ccaagtgcaa tgatcccaac attattagaa gggaagattg caatggcata 3180
ttcagaatta atgtcagtgt gtcaaagaac ttaaatttaa aattgaggcc tggagagaaa 3240
aaacctggat tttgggtgcc ccgtgtttgg gcgaatcctc ggaactttaa tttcgacaat 3300
gtgggaaacg ctatgctggc gttgtttgaa gttctctcct tgaaaggctg ggtggaagtg 3360
agagatgtta ttattcatcg tgtggggccg atccatggaa tctatattca tgtttttgta 3420
ttcctgggtt gcatgattgg actgaccctt tttgttggag tagttattgc taatttcaat 3480
gaaaacaagg ggacggcttt gctgaccgtc gatcagagaa gatgggaaga cctgaagagc 3540
cgactgaaga tcgcacagcc tcttcatctt ccgcctcgcc cggataatga tggttttaga 3600
gctaaaatgt atgacataac ccagcatcca ttttttaaga ggacaatcgc attactcgtc 3660
ctggcccagt cggtgttgct ctctgtcaag tgggacgtcg aggacccggt gaccgtacct 3720
ttggcaacaa tgtcagttgt tttcaccttc atctttgttc tggaggtaac catgaagatc 3780
atagcaatgt cgcctgctgg cttctggcaa agcagaagaa accgatacga tctcctggtg 3840

acgtcgcttg gcgttgtatg ggtggtgctt cactttgccc tcctgaatgc atatacttac 3900
atgatgggcg cttgtgtgat tgtatttagg tttttctcca tctgtggaaa acatgtaacg 3960
ctaaagatgc tcctcttgac agtggtcgtc agcatgtaca agagcttctt tatcatagta 4020
ggcatgtttc tcttgctgct gtgttacgct tttgctggag ttgttttatt tggtactgtg 4080
aaatatgggg agaatattaa caggcatgca aattttttct cggctggaaa agctattacc 4140
gtactgttcc gaattgtcac aggtgaagac tggaacaaga ttatgcatga ctgtatggta 4200
cagcctccgt tttgtactcc agatgaattt acatactggg caacagactg tggaaattat 4260
gctggggcac ttatgtattt ctgttcattt tatgtcatca ttgcctacat catgctaaat 4320
ctgcttgtag ccataattgt ggagaatttc tccttgattt attccactga ggaggaccag 4380
cttttaagtt acaatgatct tcgccacttt caaataatat ggaacatggt ggatgataaa 4440
agagaggtat tccccacgtt ccgcgtcaag ttcctgctgc ggctactgcg tgggaggctg 4500
gaggtggacc tggacaagga caagctcctg tttaagcaca tgtgctacga aatggagagg 4560
ctccacaatg gcggcgacgt caccttccat gatgtcctga gcatgctttc ataccggtcc 4620
gtggacatcc ggaagagctt gcagctggag gaactcctgg cgagggagca gctggagtac 4680
accatagagg aggaggtggc caagcagacc atccgcatgt ggctcaagaa gtgcctgaag 4740
cgcatcagag ctaaacagca gcagtcgtgc agtatcatcc acagcctgag agagagtcag 4800
cagcaagagc tgagccggtt tctgaacccg cccagcatcg agaccaccca gcccagtgag 4860
gacacgaatg ccaacagtca ggacaacagc atgcaacctg agacaagcag ccagcagcag 4920
ctcctgagcc ccacgctgtc ggatcgagga ggaagtcggc aagatgcagc cgacgcaggg 4980
aaaccccaga ggaaatttgg gcagtggcgt ctgccctcag ccccaaaacc aataagccat 5040
tcagtgtcct cagtcaactt acggtttgga ggaaggacaa ccatgaaatc tgtcgtgtgc 5100
aaaatgaacc ccatgactga cgcggcttcc tgcggttctg aagttaagaa gtggtggacc 5160
cggcagctga ctgtggagag cgacgaaagt ggggatgacc ttctggatat ttag 5214

<210> 61
<211> 1818
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 168827CB1

<400> 61
ggaaattgct tccgtgaccc tgctgcagat gggagagagg gcccattaag aagagagtgg 60
ggtcaggatc aacacacaca cttagtgtga tttaaggaaa ggaaatattt tctctttgaa 120
cttatctgga tacagtcatt ttgtctcctc ttggggatca cttgtccagc ctcaatggcc 180
tttcaggacc tcctagatca agttggaggc ctggggagat tccagatcct tcagatggtt 240
ttccttataa tgttcaacgt catagtatac catcaaactc agctggagaa cttcgcagca 300
ttcatacttg atcatcgctg ctgggttcat atactggaca atgacactat ccctgacaat 360
gaccctggga ccctcagcca ggatgccctc ctgagaatct ccatcccatt cgactcaaat 420
ctgaggccag agaagtgtcg tcgctttgtc catccccagt ggaagctcat tcatctgaat 480
gggaccttcc ccaacacgag tgagccagat acagagccct gtgtggatgg ctgggtatat 540
gaccaaagct ccttcccttc caccattgtg actaagtggg atctggtatg cgaatctcaa 600
ccactgaatt cagtagctaa atttctattc atggctggaa tgatggtggg aggcaaccta 660
tatggccatt tgtcagacag gtttgggaga aagttcgtgc tcagatggtc ttacctccag 720
ctcgccattg taggcacctg tgcggccttt gctcccacca tcctcgtata ctgctccctg 780
cgcttcttgg ctggggctgc tacatttagc atcattgtaa atactgtttt gttaattgta 840
gagtggataa ctcaccaatt ctgtgccatg gcattgacat tgacactttg tgctgctagt 900
attggacata taaccctggg aagcctggct tttgtcattc gagaccagtg catcctccag 960
ttggtgatgt ctgcaccatg ctttgtcttc tttctgttct caaggtggct ggcagagtct 1020
gctcggtggc tcattatcaa caacaaacca gaagagggct taaaggaact tacaaaagct 1080
gcacacagga atggaatgaa gaatgctgaa gacatcctaa ccatggaggt tttgaaatcc 1140
accatgaagc aagaactgga ggcagcacag aaaaagcatt ctctttgtga attgctccgc 1200
atacccaaca tatgtaaaag aatctgtttc ctgtcctttg tgagatttgc aagtaccatc 1260
ccttttgggg gccttacttt gcacctccag catctgggaa acaatgtttt cctgttgcag 1320
actctctttg gtgcagtcac cctcctggcc aattgtgttg caccttgggc actgaatcac 1380
atgagccgtc gactaagcca gatgcttctc atgttcctac tggcaacctg ccttctggcc 1440
atcatatttg tgcctcaaga aatgcagacc ctgcgtgtgg ttttggcaac cctgggtgtg 1500
ggagctgctt ctcttggcat tacctgttct actgcccaag aaaatgaact aattccttcc 1560
ataatcaggg gaagagctac tggaatcact ggaaactttg ctaatattgg gggagccctg 1620
gcttccctca tgatgatcct aagcatatat tctcgacccc tgccctggat catctatgga 1680
gtctttgcca tcctctctgg ccttgttgtc ctcctccttc ctgaaaccag gaaccagcct 1740
cttcttgaca gcatccagga tgtggaaaat gagggagtaa atagcctagc tgcccctcag 1800
aggagctctg tgctatag 1818

<210> 62

<211> 2245
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7472734CB1

<400> 62
ccctggaac agtaggatgt tggtgatgca aaagtcaatg tttaaactca acattccact 60
ttcctttaac taagaatagt ttttattaac ttttagtaaa actcagtcct agtccaaaaa 120
aagccctgct ctctgatctt tgtacaagaa catcataaag caattcactt tggattttct 180
aatatcccat ttctgagaag aatggcagac tattgaacag gtgtatttta ggtcacgtgg 240
ggactgcatc cacctgaaaa tccaccgttg acttatcagg aaactcagag atcaggatct 300
ttcacagagt agtcttttaa gaagattcag ttgtcaacag ctagcagtct ctttgccaaa 360
taattatatc tgtgacttct gaaactattt ggctgcctaa agttaaagga cttggggaaa 420
gtccctccac tgctcttctg cagtagtgtc acaccactca gtgcagggcc caccaagaag 480
aaagcagtgt caggatccac atggcactat ggtaactttg tgaaagggga cattttctcc 540
ctctgaactt ctcttcataa agtcattgtg cttcctcttg gggatcacct gttcagtctc 600
aatgggcttt gatgtgctcc tggatcaagt gggtggcatg gggagattcc agatttgtct 660
gatagctttc ttttgcatca ccaacatcct actgttccct aatattgtgt tggagaactt 720
cactgcattc acccctagtc atcgctgctg ggtccccctc ctggacaatg acactgtgtc 780
tgacaatgat accgggaccc tcagcaagga tgacctcctg agaatctcca tcccactgga 840
ctcaaacctg aggccacaga agtgtcagcg ctttatccat ccccagtggc agctccttca 900
cctgaacggg accttcccca acacaaatga gccagacacg gagccctgtg tggatggctg 960
ggtgtacgac agaagctctt tcctctccac catcgtgact gagtgggacc tggtatgtga 1020
atctcagtca ctaaaatcaa tggttcaatc cctatttatg gctgggtcac ttctgggagg 1080
tctaatatat ggccatcttt cagacaggtt tgggagaaag ttcgtgctca gatggtctta 1140
cctccagctc gccattgtag gcacctgtgc ggcctttgct cccaccatcc tcgtatactg 1200
ctccctgcgc ttcttggctg gggctgctac atttagcatc attgtaaata ctgttttgtt 1260
aattgtagag tggataactc accaattctg tgccatggca ttgacattga cactttgtgc 1320
tgctagtatt ggacatataa ccctgggaag cctggctttt gtcattcgag accagtgcat 1380
cctccagttg gtgatgtctg caccatgctt tgtcttcttt ctgttctcaa ggtggctggc 1440
agagtctgct cggtggctca ttatcaacaa caaaccagaa gagggcttaa aggaacttag 1500
aaaagctgca cacaggaatg gaatgaagaa tgctgaagac atcctaacca tggaggtttt 1560
gaaatccacc atgaagcaag aactggaggc agcacagaaa aagcattctc tttgtgaatt 1620
gctccgcata cccaacatat gtaaaagaat ctgtttcctg tcotttgtga gatttgcaag 1680
taccatccct ttttggggcc ttactttgca cctccagcat ctgggaaaca atgttttcct 1740
gttgcagact ctctttggtg cagtcaccct cctggccaat tgtgttgcac cttgggcact 1800
gaatcacatg agccgtcgac taagccagat gcttctcatg ttcctactgg caacctgcct 1860
tctggccatc atatttgtgc ctcaagaaat gcagaccctg cgtgtggttt tggcaaccct 1920
gggtgtggga gctgcttctc ttggcattac ctgttctact gcccaagaaa atgaactaat 1980
tccttccata atcaggggaa gagctactgg aatcactgga aactttgcta atattggggg 2040
agccctggct tccctcatga tgatcctaag catatattct cgaccgctgc cctggatcat 2100
ctatggagtc tttgccatcc tctctggcct tgttgtcctc ctccttcctg aaaccaggaa 2160
ccagcctctt cttgacagca tccaggatgt ggaaaatgag ggagtaaata gcctagctgc 2220
ccctcagagg agctctgtgc tatag 2245

<210> 63
<211> 3196
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7473473CB1

<400> 63
gcggcggccg ggggagcgct actaccatga actgcctggt cctcctcccc agagctgctc 60
atccgggtcg ggctggagac acagtcaggg gaccccgtcg ccgccgccgc gccccctctt 120
ctttcggctc aatcttctct tccacctttt cctcctcttc ctccaccttc tttgcctgca 180
tccccccctc ccccgccgcg gatcctggcc gctgctctcc agacccagga tgccgggggg 240
caagagaggg ctggtggcac cgcagaacac atttttggag aacatcgtca ggcgctccag 300
tgaatcaagt ttcttactgg gaaatgccca gattgtggat tggcctgtag tttatagtaa 360
tgacggtttt tgtaaactct ctggatatca tcgagctgac gtcatgcaga aaagcagcac 420
ttgcagtttt atgtatgggg aattgactga caagaagacc attgagaaag tcaggcaaac 480
ttttgacaac tacgaatcaa actgctttga agttcttctg tacaagaaaa acagaacccc 540

tgtttggttt tatatgcaaa ttgcaccaat aagaaatgaa catgaaaagg tggtcttgtt 600
cctgtgtact ttcaaggata ttacgttgtt caaacagcca atagaggatg attcaacaaa 660
aggttggacg aaatttgccc gattgtacacg ggctttgaca aatagccgaa gtgttttgca 720
gcagctcacg ccaatgaata aaacagaggt ggtccataaa cattcaagac tagctgaagt 780
tcttcagctg ggatcagata tcctttcctca gtataaacaa gaagcgccaa agacgccacc 840
acacattatt ttacattatt gtgcttttaa aactacttgg gattgggtga ttttaattct 900
taccttctac accgccatta tggttcctta taatgtttcc ttcaaaacaa agcagaacaa 960
catagcctgg ctggtactgg atagtgtggt ggacgttatt tttctggttg acatcgtttt 1020
aaattttcac acgactttcg tggggcccgg tggagaggtc atttctgacc ctaagctcat 1080
aaggatgaac tatctgaaaa cttggtttgt gatcgatctg ctgtcttgtt taccttatga 1140
catcatcaat gcctttgaaa atgtggatga gggaatcagc agtctcttca gttctttaaa 1200
agtggtgcgt ctcttacgac tgggccgtgt ggctaggaaa ctggaccatt acctagaata 1260
tggagcagca gtcctcgtgc tcctggtgtg tgtgtttgga ctggtggccc actggctggc 1320
ctgcatatgg tatagcatcg gagactacga ggtcattgat gaagtcacta acaccatcca 1380
aatagacagt tggctctacc agctggcttt gagcattggg actccatatc gctacaatac 1440
cagtgctggg atatgggaag gaggacccag caaggattca ttgtacgtgt cctctctcta 1500
ctttaccatg acaagcctta caaccatagg atttggaaac atagctccta ccacagatgt 1560
ggagaagatg ttttcggtgg ctatgatgat ggttggcgct cttctttatg caactatttt 1620
tggaaatgtt acaacaattt tccagcaaat gtatgccaac accaaccgat accatgagat 1680
gctgaataat gtacgggact tcctaaaact ctatcaggtc ccaaaaggcc ttagtgagcg 1740
agtcatggat tatattgtct caacatggtc catgtcaaaa ggcattgata cagaaaaggt 1800
cctctccatc tgtcccaagg acatgagagc tgatatctgt gttcatctaa accggaaggt 1860
ttttaatgaa catcctgctt ttcgattggc cagcgatggg tgtctgcgcg ccttggcggt 1920
agagttccaa accattcact gtgctcccgg ggacctcatt taccatgctg gagaaagtgt 1980
ggatgccctc tgctttgtgg tgtcaggatc cttggaagtc atccaggatg atgaggtggt 2040
ggctatttta gggaagggtg atgtatttgg agacatcttc tggaaggaaa ccacccttgc 2100
ccatgcatgt gcgaacgtcc gggcactgac gtactgtgac ctacacatca tcaagcggga 2160
agccttgctc aaagtcctgg acttttatac agcttttgca aactcctttt caaggaatct 2220
cactcttact tgcaatctga ggaaacggat catctttcgt aagatcagtg atgtgaagaa 2280
agaggaggag gagcgcctcc ggcagaagaa tgaggtgacc ctcagcattc ccgtggacca 2340
cccagtcaga aagctcttcc agaagttcaa gcagcagaag gagctgcgga atcagggctc 2400
aacacagggt gaccctgaga ggaaccaact ccaggtagag agccgctcct tacagaatgg 2460
agcctccatc accggaacca gcgtggtgac tgtgtcacag attactccca ttcagacgtc 2520
tctggcctat gtgaaaacca gtgaatccct taagcagaac aaccgtgatg ccatggaact 2580
caagcccaac ggcggtgctg accaaaaatg tctcaaagtc aacagcccaa taagaatgaa 2640
gaatggaaat ggaaaagggt ggctgcgact caagaataat atgggagccc atgaggagaa 2700
aaaggaagac tggaataatg tcactaaagc tgagtcaatg gggctattgt ctgaggaccc 2760
caagagcagt gattcagaga acagtgtgac caaaaaccca ctaagaaaaa cagattcttg 2820
tgacagtgga attacaaaaa gtgaccttcg tttggataag gctggggagg cccgaagtcc 2880
gctagagcac agtcccatcc aggctgatgc caagcacccc ttttatccca tccccgagca 2940
ggccttacag accacactgc aggaagtcaa acacgaactc aaagaggaca tccagctgct 3000
cagctgcaga atgactgccc tagaaaagca ggtggcagaa attttaaaaa tactgtcgga 3060
aaaaagcgta ccccaggcct catctcccaa atcccaaatg ccactccaag tacccccccа 3120
gataccatgt caggatattt ttagtgtctc aaggcctgaa tcacctgaat ctgacaaaga 3180
tgaaatccac ttttaa 3196

<210> 64
<211> 1602
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 7477725CB1

<400> 64
atggcctttg aggagctctt gagtcaagtt ggaggccttg ggagatttca gatgcttcat 60
ctggttttta ttcttccctc tctcatgtta ttaatccctc atatactgct agagaacttt 120
gctgcagcca ttcctggtca tcgttgctgg gtccacatgc tggacaataa tactggatct 180
ggtaatgaaa ctggaatcct cagtgaagat gccctcttga gaatctctat cccactagac 240
tcaaatctga ggccagagaa gtgtcgtcgc tttgtccatc cccagtggca gcttcttcac 300
ctgaatggga ctatccacag cacaagtgag gcagacacag aaccctgtgt ggatggctgg 360
gtatatgatc aaagctactt cccttcgacc attgtgacta agtgggacct ggtatgtgat 420
tatcagtcac tgaaatcagt ggttcaattc ctacttctga ctggaatgct ggtgggaggc 480
atcataggtg gccatgtctc agacaggttt gggcgaagat ttattctcag atggtgtttg 540
ctccagcttg ccattactga cacctgcgct gccttcgctc ccaccttccc tgtttactgt 600
gtactacgct tcttggcagg tttttcttcc atgatcatta tatcaaataa ttctttgccc 660

attactgagt ggataaggcc caactctaaa gccctggtag taatattgtc atctggtgcc 720
cttagtattg gacagataat cctgggaggc ttggcttatg tcttccgaga ctggcaaacc 780
ctgcacgtgg tggcgtctgt acctttcttt gtcttctttc ttctttcaag gtggctggtg 840
gaatctgctc ggtggttgat aatcaccaat aaactagatg agggcttaaa ggcacttaga 900
aaagttgcac gcacaaatgg aataaagaat gctgaagaaa ccctgaacat agaggttgta 960
agatccacca tgcaggagga gctggatgca gcacagacca aaactactgt gtgtgacttg 1020
ttccgcaacc ccagtatgcg taaaaggatc tgtatcctgg tatttttgag atttgcaaac 1080
acaatacctt tttatggtac catggtcaat cttcagcatg tggggagcaa cattttcctg 1140
ttgcaggtac tttatggagc tgtcgctctc atagttcgat gtcttgctct tttgacacta 1200
aatcatatgg gccgtcgaat aagccagata ttgttcatgt tcctggtggg cctttccatt 1260
ttggccaaca cgtttgtgcc caaagaaatg cagaccctgc gtgtggcttt ggcatgtctg 1320
ggaatcggct gttctgctgc tacttttttcc agtgttgctg ttcacttcat tgaactcatc 1380
cccactgttc tcagggcaag agcttcagga atagatttaa cggctagtag gattggagca 1440
gcactggctc ccctcttgat gaccttaacg gtatttttta ccactttgcc atggatcatt 1500
tatggaatct tccccatcat tggtggcctt attgtcttcc tcctaccaga aaccaagaat 1560
ctgcctttgc ctgacaccat caaggatgtg gaaaatcagt ga 1602